

City University of New York (CUNY)

CUNY Academic Works

Dissertations, Theses, and Capstone Projects

CUNY Graduate Center

2-2015

What the Tortoise Said to Kripke: the Adoption Problem and the Epistemology of Logic

Romina Padro

Graduate Center, City University of New York

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/gc_etds/603

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).

Contact: AcademicWorks@cuny.edu

WHAT THE TORTOISE SAID TO KRIPKE:
THE ADOPTION PROBLEM AND THE EPISTEMOLOGY OF LOGIC

by

ROMINA PADRO

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the
requirements for the degree of Doctor of Philosophy,
The City University of New York

2015

© 2015

ROMINA PADRO

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Gary Ostertag

October 7th, 2014

Date

Chair of Examining Committee

Rosemarie Iannuzzi

October 7th, 2014

Date

Executive Officer

Michael Devitt, Advisor

Arnold Koslow

Michael Levin

Paul Boghossian

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

WHAT THE TORTOISE SAID TO KRIPKE:

THE ADOPTION PROBLEM AND THE EPISTEMOLOGY OF LOGIC

by

ROMINA PADRO

Advisor: Michael Devitt

In chapter 1 I introduce the main topics to be addressed and provide a summary of the dissertation. In chapter 2 I summarize Lewis Carroll's famous note "What the Tortoise Said to Achilles" and briefly review some of its most influential interpretations. The rest of the chapter is devoted to Kripke's unpublished interpretation of Carroll's note and the moral he draws from it: in section 2.2 I present and discuss what I call the "adoption problem" and in section 2.3 I clarify certain aspects of it. In chapter 3 I consider a modification of the original set up of the adoption problem in terms of rules of inference and argue that they cannot help us overcome it. I also compare Kripke's interpretation of Carroll with his famous interpretation of Wittgenstein in *Wittgenstein on Rules and Private Language* and discuss some points of connection between Wittgenstein's and Kripke's views on the relation between inferential acts and logical principles or rules.

The following three chapters are meant to discuss, in view of the adoption problem, versions of what have perhaps been the most influential proposals for the justification of basic logical principles: empiricism *a la* Quine, rational intuition, and concept-constitution. In chapter 4 I examine Kripke's use of the adoption problem to criticize Quine's views in "Two Dogmas of Empiricism." Since Kripke argues that Quine fails to see that the same objections he directs against Carnap in "Truth by Convention" can be applied to his own point of view in "Two

Dogmas,” I start with a brief review of Carnap’s proposal and Quine’s famous objection to it. I continue with Quine’s “Two Dogmas” views, examine Kripke’s criticisms, and claim that Quine’s argument against Carnap should not be identified with Kripke’s argument against Quine. I also distinguish what I call the “standard objection” to Quine’s views from the objection based on the adoption problem and critically consider an attempt to overcome the latter by introducing some modifications to the Quinean system. In the final section I discuss Kripke’s Carroll-inspired metaphysical argument against the Cartesian doctrine on the creation of eternal truths and propose a related epistemological version of the argument. In chapter 5 I introduce Kripke’s intuition-based proposal for accounting for basic logical beliefs and inferences outlined in his talks and seminars and argue that it should be rejected because it overlooks his own adoption problem. I also note some possible variations of the proposal and claim that they do not fare better. In chapter 6 I focus on Boghossian’s concept-constituting account. I discuss in detail his proposal for dealing with the justification of basic logical beliefs and some of the criticisms it has received. I argue that the main problem of this proposal is that it fails to accommodate the lessons that can be drawn from the adoption problem concerning the nature of inferring.

In chapter 7 I reconsider the significance of the adoption problem and focus on Ryle’s famous distinction between *knowing how* and *knowing that*. In section 7.2 I argue that “intellectualism,” the claim that *knowledge how* can be defined in terms of propositional knowledge (*knowledge that*), is not an option when it comes to basic deductive inferences and that some sort of *knowing how* is needed for an account of both propositional knowledge of basic logical principles and inferential uses (Kripke’s interpretation of Carroll will play a key role in my argument). I then discuss two different ways of understanding the nature of basic inferences: a view in which deductive inferential rules are ‘hardwired’ in us and govern our inferences without being represented, and a

view where basic inferential acts are taken as a primitive ability, not internally informed by logical principles, though encouraged and improved by our direct immersion in inferential practices. I discuss some problems affecting the former proposal problems and leave the prospects of the latter as an open question. Finally, I argue that independently of what it is ultimately to be made of the nature of basic inferences, the adoption problem already indicates the need for rethinking the problem of the justification of basic logical principles and rules and what we hoped to achieve with its solution.

To the Memory of Jonathan E. Adler

Acknowledgements

The late Jonathan Adler was my advisor when the first three quarters of this dissertation were being prepared. He was extremely supportive and helped me shape the main ideas I defend here. Perhaps one of the most difficult parts of the writing process was to set aside some issues I originally thought should be included. Jonathan had a crucial role in making me see that not everything I wished to include could possibly be discussed. Even in the last months of his life he continued to be there for me, returning comments and checking on my progress. I feel very fortunate to have had him as my advisor and will never forget his kindness and intelligence.

Michael Devitt stepped up as my advisor when Jonathan passed away. He gave me encouragement at key moments, always-needed pressure, and valuable and abundant criticism and advice. His die-hard Quinean heart has challenged me and pushed me to better articulate my views. We have certainly disagreed on several issues; I won some battles, he won others, but I know these issues are far from settled and that we will continue discussing them in the future. Working with him on this project and at the Saul Kripke Center has been a lot of fun and I am profoundly indebted to him for his patience and friendship.

My committee members and readers have been very generous with their time and have offered thoughtful criticism and advice. Arnold Koslow has been enormously supportive of me and my work since we first met, and I am grateful to him for his vote of confidence and his valuable suggestions. Michael Levin has been a very careful and dedicated reader, making many helpful remarks and criticisms. His interest in my project and the discussions I had with him over the years have been very important to me. Paul Boghossian kindly agreed to set time aside to be one of my readers during a sabbatical. His excellent criticisms will surely keep me occupied in the years to come, and I hope he is aware of how much his positive comments on my writing have meant to me. Gary Ostertag read a number of versions of this dissertation, anticipated different criticisms, and made suggestions to improve my arguments. His generosity, friendship, and support has been vital to me during these years. Each of the members of my committee contributed in different ways to make my work better. And for that, I am truly grateful.

Meeting Saul Kripke was, without a doubt, the best part of studying at the Graduate Center. It was, indeed, a life-changing event. I have worked with him at the Saul Kripke Center for many

years now and the depth of his intelligence and knowledge on virtually any topic never ceases to amaze me. But it is his contagious enthusiasm for philosophical problems – which has served as a reminder, in more than one occasion, of my true reasons for pursuing this subject – and his unwavering intellectual honesty, that I admire the most. It is hard for me to express how grateful I am for getting to know Saul and for having the privilege of working with such an exceptional person. It will be clear to the reader that I have benefited tremendously from conversations with him on the nature of logic and on other philosophical topics. Indeed, I doubt there is a page in this dissertation that doesn't betray his influence on my work. I also want to thank him for reading and commenting on different versions, for letting me quote his unpublished work, and for agreeing to write a summary of his argument against Descartes' views on the eternal truths to be included here (chapter 4).

Versions of the first three chapters of this dissertation were presented at Fudan University, Pekin University, Southwest University at Chongqing, Shandong University, and the University of Nebraska at Omaha. Parts of the last four chapters were presented at a workshop on the epistemology of logic in Argentina and to the SADAF (Argentina) logic seminar. I would like to thank my commentators and the audiences at these events for their helpful questions and observations.

Writing can sometimes be a lonely process. I was lucky to have three main writing buddies: Rosemarie Twomey, Ignacio Moreno, and Sirius. The many meetings with Rosie at the New York Public Library made me look forward to spending hours there, in spite of the fact that it was often a sunny Saturday or Sunday afternoon. Ignacio, my oldest nephew, also shared a number of writing hours with me. Conversations with both of them during writing breaks tended to be memorable and I have enjoyed, perhaps a bit too much, every one of them. There were no chats with Sirius, my mini schnauzer. But he spent many nights sleeping on my feet, probably wondering why we weren't going to bed already. I am grateful for his silence and for the warm feet.

Other members of my family also had important roles during these years. Lala, my grandmother, passed away in the beginning of my PhD, but somehow managed to remain very present to me. The memory of her energy and enthusiasm continues to be contagious. Silvina, my sister, undertook the tiresome task of checking my writing progress. Though I am grateful for

those calls, I can't say that I will be missing them! Matías, my youngest nephew, supplied authoritative information on children's movies and a much-needed example. Laura and Natalia were, as always, there for me when I needed advice, encouragement, or just to complain about lack of progress. My Argentinian philosophical family, the GAF group, also deserve a special acknowledgement. In particular, I would like to thank Eduardo Barrio and Alberto Moretti, who were my advisors and continued accompanying me through the years.

Phrases such as "this work would not have been possible without the support of my partner" are hardly uncommon when it comes to giving thanks. This is not an exception. Here, however, the meaning should be taken literally. It was my husband, Fernando Birman, who persuaded me to come to New York to work on a PhD, and it was him again who persuaded me (and made sure) that I finish it. Thank you. I hope some day I would have the chance to retrieve in kind and do for you what you have done for me.

Table of Contents

Chapter 1. Introduction	1
1.1 Basic Logical Principles and Their Justification	1
1.2 Setting Justifications Aside	8
1.3 The Carroll-Kripke Diagnosis	17
1.4 Brief Outline	18
Chapter 2. Carroll's Regress and the Adoption Problem	21
2.1 Carroll's Note	22
2.2 The Adoption Problem	31
2.2.1 Harry and the Tortoise	35
2.2.2 Some Clarifications	41
Chapter 3: Rules, Examples, and the Wittgenstein	
Connection	52
3.1 Basic Inferential Rules	52
3.2 Examples: The Last Attempt	63
3.3 The Wittgenstein Connection	66
3.3.1 <i>WRPL</i> and the Adoption Problem	67
3.3.2 Rethinking the Connection	72
3.3.3 Kripke and Wittgenstein on Rules	
and Inferences	79

Chapter 4: Quine and Descartes on the Status of Logic:	
Kripke's Critique	86
4.1.1 Quine's version of the Carroll Argument and the	
"Two Dogmas" Conception of Logic	86
4.1.2 The Standard Objection	98
4.1.3 "Two Dogmas" and the Adoption Problem	106
4.2 Descartes on Eternal Truths	122
Chapter 5: Kripke on Logic and Intuition	131
5.1 Self-Evidence and the Analogy between Formal	
Systems and Non-Euclidean Geometries	132
5.2 The Role of Intuitions and the Adoption Problem	150
Chapter 6: Concept-Constituting Accounts	158
6.1 Clearing the Ground for a Circular Justification	162
6.2 Vindicating Rule-Circularity	168
6.3 Warrant Transfer by Inference	
and Concept-Constitution	174
6.4 Boghossian's Proposal and the Adoption Problem	186
Chapter 7: On the Nature of Basic Inferences	194
7.1 Ryle on Knowing How / Knowing That	196
7.2.1 Basic Deductive Inferences as Knowing How	209
7.2.2 Hardwired Rules of Inference	212
7.2.3 Basic Inferring Without Rules	222

CHAPTER 1

Introduction

1.1 Basic Logical Principles and Their Justification

Consider the following: I tell my seven year-old nephew, Matías, that if Sunday is a sunny day, I will take him to the zoo. On Sunday he wakes up at 6 a.m., checks the sky (which turns out to be unmistakably sunny) and proceeds to wake me up demanding to be taken to the zoo. Matías inferred in accordance with *Modus Ponendo Ponens* (MPP), a very familiar pattern of inference (i.e. if Sunday is a sunny day, Romina will take me to the zoo. Sunday is a sunny day, so Romina will take me to the zoo).

While at the zoo, he informs me that all the animals in the movie *Madagascar* talk. I ask: “Does Alex the lion talk?” (It’s too early for me and I’m still slightly asleep.) Impatiently, he replies “yes! I told you that *all* the animals in the movie talk!” He has now faulted me for not inferring according to Universal Instantiation (UI) (i.e. all the animals in the movie *Madagascar* talk; Alex, the lion in *Madagascar*, talks).

Children (and arguably some animals)¹ are able to engage in inferences that accord with basic logical principles – that is, logical principles that are not derived from other logical principles.²

¹ The issue of whether animals are capable of deductive inference remains unresolved. The debate began in antiquity, most famously with Aristotle’s claim that only humans can reason and Sextus Empiricus’s example of Chrysippus’s dog. Descartes’s fierce opposition and Hume’s defense are also classic stances in the debate. Disagreements among philosophers continue to be common, with Davidson (1982) among the most prominent opponents of animal reasoning and Fodor (1987) and Carruthers (2004) among its supporters. For a history and summary of the different aspects of the debate, see the introduction to Lurz (2009), Wilson (1995), and Sorabji (1993) (my thanks to Jonathan Adler for these references). More attention, it seems to me, ought to be paid to research on animal behavior (especially by some of those

We, as adults, seem to rely on such inferences with total confidence. Yet, if someone were obnoxious enough to ask *why* we are so confident, we would be in serious trouble.

Suppose we are at a party talking to a group of friends and say: “If Maria comes, she will sing for us.” One friend notes: “She’s already here,” and we immediately add: “That’s great, I was looking forward to listening to her. We’d better sit next to the piano.” Someone overhears the conversation and asks why we think Maria will sing. Our first reply is likely to consist in a repetition of the premises, maybe adding some reasons for them, such as “I talked to Maria earlier today and she said she would definitely sing if she managed to come” and “she *did* come, she’s right there by the bar.”

Such an answer would be sufficient in nearly all cases. But imagine that our questioner remains unmoved: he assents to the premises and our reasons for believing them but keeps asking why we believe Maria will sing. Eventually, it becomes clear that he is questioning the inference itself, whether our acceptance of the premises gives us reason for the conclusion. Most people (the wise ones) will cut and run. Some, in particular those that still hold on to a few memories from Logic 101, may be inclined to start talking about validity: since MPP inferences ($p \rightarrow q$, p/q) are valid, it is impossible for the premises to be true and the conclusion false; our inference happens

who oppose it, as Davidson himself), though of course clarification as to what is meant by “capable of deductive inference” is also needed (but this applies to the human case as well).

² UI and MPP are usually considered basic in this sense, but logical systems may pick different principles as basic, and a propositional system would, of course, lack UI. Clearly, not all principles can be derived from other principles, some have to count as basic. I will just assume here that MPP and UI are basic (see fn. 56, for some discussion of this issue) and valid, truth-preserving whenever the premises are true (non-classical logicians who disagree with this last assumption can add ‘classically’ before ‘valid,’ disagreements of this kind won’t be our main concern).

Also, as we go on, we will need to stress the distinction between logical implications, associated conditionals, and rules of inference, as well as the connections among these notions. For now, I’ll just talk of ‘logical principles’ (the importance of keeping these distinctions in place will be discussed in chapter 2).

Finally, formal systems won’t be my main concern. Instead, I will focus on basic logical principles we can be said to infer in accordance with, like the ones given by the examples above.

to be an instance of MPP, so we are justified in believing that Maria will sing (assuming the truth of the premises).

But if our questioner has been willing to go as far as this, he may be willing to go a bit further. He may want to ask how we know that MPP is a valid inference, and may even ask how this validity is supposed to vindicate our own inference. We did, if he is not mistaken, seem to be using MPP again when we argued from the validity of MPP to the conclusion that we are justified in believing that Maria will sing.

By then we have probably grown impatient with our questioner – ‘skeptic,’ we now realize, is a more appropriate description – and may decide to simply point to the piano. Maria *is* standing next to it with a microphone in her hand, and the first notes can already be heard.

“Surely, just a coincidence,” he says. Luckily for us, we now have a way to end the conversation, if not by persuading him, at least by pointing out that it would be rude to keep talking while Maria sings.

Is an epistemic justification of basic logical principles needed? Is it possible? Answering ‘No’ to the second question tends to encourage a negative answer to the first one, and indeed many philosophers have taken this view.³ The problem is that if MPP and UI are to count as basic principles, we cannot appeal to other logical principles to justify them. Semantics doesn’t get us

³ Dummett claims that this is the ‘tacit’ position of most philosophers (1973b: 292; see also 1991: 23). Surprisingly, he mentions Goodman but not Frege, who appears to be a clear case of this attitude. For Frege, “the question why and with what right we acknowledge a law of logic to be true, logic can answer only by reducing it to another law of logic. Where that is not possible, logic can give no answers.” (1893: 204) But he showed little interest in looking for the answer somewhere else. Frege does, however, give reasons for acknowledging them: we must do so “if we are not to throw our thought into confusion and in the end renounce judgment altogether.” (p. 204) For him, the discovery of beings that contradict the laws of logic would be the discovery of “a hitherto unknown kind of madness.” (p. 203) The laws of logic are the laws of truth, and as such they are universal, objective, and self-evident, “boundary stones set in an eternal foundation, which our thought can overflow but not dislodge.” (p. 203)

What I say in the text above is not an attempt to describe a specific view, but a general attitude towards the problem.

very far either. For MPP, we could *argue*: a conditional is true if the consequent is true whenever the antecedent is true. So if you know that a certain conditional is true and that its antecedent is true, it follows that the consequent is true. Yet, this appeal to the material conditional's truth-table uses MPP, the very same principle we are trying to establish. Similar problems arise for other basic logical principles like UI: for a universal statement to be true each instance in its domain has to be true. Thus, in the case of any particular universal statement, if it is true, it follows that any instance in its domain is true. Once again, we are trying to justify UI here by appealing to UI itself. Intuitively, the problem is that attempts to establish the validity of these principles will either require the use of those basic principles or other principles that are derived from them. But then the arguments are circular.

These issues led many to either give up or simply ignore the whole project of seeking a justification of basic principles. And though this attitude might seem like a mere shrug of the shoulders, some remarks can quickly be added to dissipate our justification anxieties. We can always count on Wittgenstein to remind us that, like it or not, justifications must “come to an end somewhere” (see *Philosophical Investigations* – henceforth *PI*: §1, § 217, and *On Certainty*: § 204). MPP and UI are supposed to be basic and it is hardly surprising we have reached bedrock with them.⁴ We should then accept our own limits and deal with the fact that we cannot start from nowhere, some things need to be granted and basic principles are obvious candidates. This attitude, we may want to add, far from being irresponsible, would be a sign of intellectual maturity, a sign of having outgrown an overambitious justificational stage that is neither needed nor possible and moved on to more sensible enterprises. The moral of our encounter with the questioner should

⁴ It seems to me, however, that this is not a use of the justifications phrase and the “bedrock” metaphor (*PI*: §217) that Wittgenstein would approve. For discussion of his views on the relation between inferences and logical principles, see chapter 3, sections 3.3.2 and 3.3.3.

hence be that there is no need to engage with (much less be!) the party's annoying questioner. Walking away is the right thing to do.

Others,⁵ however, think that it is too early for capitulation. Going along with the impossibility of giving a justification of basic logical principles, we are warned, opens the door to an abrasive form of skepticism. Basic logical principles have traditionally counted as one of the most solid parts of our knowledge, and acknowledging that nothing can be said to support our claims to this knowledge would leave us in an awkward position.

But there are also further reasons for concern. As our party exchange suggests, when an inference we make is challenged, trying to ground it on an appeal to the good standing of the principle or principles involved (assuming, of course, that we possess the relevant knowledge), presents itself as the obvious option. Though we don't usually find ourselves in such a situation, an appeal to the validity of the principles might be quite natural, even when no skeptical scenario is at play – say, because someone is disputing a conclusion or we are evaluating someone else's inferences. Why do we bother? We seem to think that we are adding support to the inferences by doing so, but it is hard to see what's the point when we can give no grounds for thinking that validity claims will do any better than our inferences themselves. (Better keep 101 Logic students in the dark if this is the case: it certainly seems to back the typical “this course is useless” remark.)

The core worry behind this is that we may not be entitled⁶ to use inferences that accord with basic principles if we cannot give a justification for the validity of the principles themselves.⁷

⁵ See, for example, Boghossian (2000, 2001a, 2003), Peacocke (1998a, 2004, 2008), Wright (2001, 2004b), Sher (2011, 2013), Schechter and Enoch (2006 and 2008), and Schechter (2013).

⁶ I follow Boghossian (1996: 223 and 2000: 230, fn. 3) in reserving the term ‘entitlement’ for the *use* of inferences that accord with basic principles. He, in turn, is following Burge's well-known distinction between justification and entitlement. Entitlements are “epistemic rights or warrants that need not be understood or even be accessible to the subject” (Burge 1993: 458); justifications, on the other hand, “must be available in the cognitive repertoire of the subject”, they “involve reasons that people have and

Needless to say, if this worry were genuine, we would be in pretty bad shape: not only basic inferential uses but also the beliefs we may draw as their conclusions would be in shaky territory. And considering that the inferences in question are the most basic ones, the list of suspect beliefs will surely be extremely long. Those who have ignored or given up justification attempts at least owe us, from this point of view, a plausible explanation of why we are entitled to infer in accordance with basic principles when no justification is forthcoming.

Take Matías, for instance. We feel that he is as warranted as he can be in holding that I will take him to the zoo, and that no better reason than his own use of MPP can be given for it (i.e. if Sunday is a sunny day, Romina will take me to the zoo. Sunday is a sunny day, so Romina will take me to the zoo). If it turns out that we are unable to justify MPP, would that also mean that Matías's *use* of an inference in accordance with MPP is unwarranted? We might then want to ask what is going on here: how is it that he has what we take to be as good a reason as there can be for believing that I will take him to the zoo and his use of the inference may not be warranted at the same time?

It is also possible that we are worrying over nothing. After all, what we do when we reason may have nothing or little to do with logical principles. Harman has in fact been insisting on this

have access to" (p. 459). So, people can be entitled even if they are unable to provide or consider a possible justification.

⁷ As Boghossian puts it, "if it is impossible for the claim that MPP is truth-preserving to be justified that would seem to imply that our use of MPP is also unjustifiable. For how could we be entitled to reason according to a given inference rule if it is impossible for the claim that that rule is truth-preserving to be justified in the slightest?" (2000: 234). Boghossian thinks that assuming the impossibility of giving a justification of basic inferences leads to an unstable global relativism (see 2001a: 246ff). Wright expresses a similar worry when he says that to claim that we cannot justify our basic inferences "would be to say that, at the most fundamental level of our reasoning, we hold ourselves subject to an objective and rational constraint – that of truth preservation – while not having, or at any rate not being able to produce the slightest reason for supposing that our basic inferential norms measure up to it." (2001a: 43)

point in a series of writings dating back to the 70s.⁸ Reasoning, he thinks, should be sharply distinguished from argument or proof. While reasoning as “a reasoned change in view” is a procedure for finding out what we ought to believe and, as such, it is constrained by our human limitations and finite powers, argument or proof “in accordance with the rules of logic” (1984: 107) is the process of drawing “a conclusion from premises via a series of intermediate steps.” (1986: 3)

Logic is one thing, the theory of reasoning quite another: logic’s main interest lies in the notion of logical consequence or validity, it seeks to establish logical relations among propositions or statements and is blind to considerations regarding the maximal preservation and coherence of our overall corpus of beliefs, which, in contrast, are central for a theory of rational change in view. Some of the grounds Harman gives for keeping this distinction straight are, for example, that when we reason, we do not – and should not – believe all the logical consequences of our beliefs, and that logical entailment need not indicate that a conclusion should be believed – if we find it implausible (say, in conflict with other beliefs), we may well be advised to revise our premises. Moreover, whereas adding more premises to a logically valid inference won’t affect its validity, having more information may lead a rational agent to reject a conclusion she had validly drawn.⁹

Following Harman we could use these and similar considerations to further dispute that subjects reason deductively at all (1986: 6), challenging the connection between the normative role of logic and reasoning. (In Harman’s own words: “there seems to be no significant way in

⁸ See Harman (1986: ch. 1 & 2). Also, Harman (1973, 1984, 1995, 2002, 2009), and Harman and Kulkarni (2007).

⁹ Reasoning, then, violates the monotonic property of standard logical systems, where if $\Gamma \models \phi$ and $\Gamma \subseteq \Delta$ then $\Delta \models \phi$.

which *logic* might be *specially* relevant to reasoning” (2009: 334, his emphasis).) Our worries, we could then claim, were misplaced: we need not be concerned about where our entitlement to employ logical principles could possibly come from. The whole point is that we don’t employ them at all.¹⁰

But before going on with this, another note of caution should be added. *Pace* Harman’s efforts, his view has not gained many followers. Indeed, most writers take for granted that we reason deductively and simply assume that there is a normative connection between logical principles and reasoning.¹¹ Unfair to Harman as this might be, we do seem sometimes to conform to logical principles when we infer. And, though this is far from conclusive, we shouldn’t overlook the fact that since its very Aristotelian beginnings logic has been conceived as having a normative role. (For more on Harman’s views see pp. 15 ff below and chapter 7.)

1.2 Setting Justifications Aside

Crispin Wright likes to point out that the epistemology of logic gives us a set of different but interrelated problems.¹² They can be characterized as follows:

¹⁰ See (2009: 334), where Harman disputes whether we “employ a logic” when we reason.

¹¹ MacFarlane (2004) and Field (2009) are important exceptions. They discuss Harman’s arguments and propose different bridge principles for a normative connection between reasoning and logic. See also Sainsbury (2002: especially pp. 1-4), where he defends the “somewhat naïve” opinion that, “Logic ought to guide our thinking. It is better, more rational, more intelligent to think logically than to think illogically. Illogical thought leads to bad judgment and error. In any case, if logic had no role to play as a guide to thought, why should we bother with it?” (Sainsbury 2002: 1). Frege was very emphatic on this point. He says: “Anyone who has recognized a law of truth [for him the laws of logic are laws of truth] has thereby also recognized a law that prescribes how judgments should be made, wherever, whenever and by whomever they may be made.” (1893: 204).

¹² My characterization of the problems is somewhat different from Wright’s. He says: “Logic gives rise to three hard, interrelated *epistemological* problems – two of them (at least until quite recently) surprisingly neglected. One is the problem of saying what exactly, at the level of a movement of thought, an inference is – what it is for a belief to be formed as a result of inference [my (c)]. The second is to determine under

(a). *The justification problem*: we take ourselves to *know* that our basic logical principles are correct. Is this so? And if it is, how is this knowledge to be justified?

(b). *The entitlement problem*: in figuring out what to believe, we perform inferences that accord with basic logical principles and seem to assume that whatever justification we have for the premises of for example a MPP or UI inference passes on to the conclusions we have come to believe; how is our entitlement to the use of those inferences and the corresponding idea about the transfer of justification be accounted for?

(c). And finally, *the inference problem*, the problem concerning the nature of basic deductive inferences: what is it for a thinker to accept or form a belief as the result of a (basic) inference? Under what conditions is it possible?

As Wright (2004b: 155) notes, most efforts have focused on the justification problem.¹³ The choice is not surprising: it is reasonable to expect that a satisfactory answer to (a) would take care

what circumstances an inference transmits a thinker's justification for its premisses to its conclusion – under what conditions a thinker is in position genuinely to learn by a particular inference [my (b)]. But it is the third that has attracted most attention [...] We know, we think, that *modus ponens*, for instance, is a valid rule, and that this knowledge is as rock-solid as any we have. But it is notoriously difficult to say anything satisfactory about the provenance of such knowledge: about how it is, or might in principle be achieved [my (a)].” (2004b: 155. My emphasis.) See also the beginning of Wright (2014).

¹³ There are multiple examples of this. As an illustration, consider the following remark by Gillian Russell: “The central question in the epistemology of logic is how we come to be justified in believing claims which attribute one of the logical properties—logical truth, logical consequence, equivalence, unsatisfiability etc.—to linguistic items like sentences and schemas.” (2014: 161)

of (b), giving the grounds needed for inferential uses.¹⁴ But this connection has been mostly assumed, and (b) and (c) have not usually received much attention.

Concentrating on (a) has not proved very rewarding. We have already discussed the circularity problem we encounter when trying to argue in support of basic logical principles. Indeed, if circularity is to be avoided,¹⁵ it looks as if we need to find a way of justifying our knowledge of basic logical principles that doesn't itself require the use of inferences in accordance with those principles – a direct, non-inferential approach.

And we have additional baggage to consider in dealing with (a) since logic has been historically thought to be a paradigmatic case of *a priori* knowledge – i.e. knowledge, roughly stated, that is independent from our experiences of the world.¹⁶ As one would expect, there is nothing light about this baggage. For centuries the *a priori* has been the defining battleground of the dispute between rationalists and empiricists, and we have little reason for thinking that the debate will be resolved any time soon.¹⁷

¹⁴ Supposing, *contra* Harman, that there is a normative connection between reasoning and logical principles.

¹⁵ Boghossian has developed an influential alternative that claims that it need not be (we will discuss it in chapter 6). See also Dummett (1973b: 296), where he draws a distinction between a “suasive” argument, which intends to persuade someone of its conclusion, and an “explanatory” argument, which intends to explain the conclusion to someone who already accepts it. He goes on to argue that in the case of the justification of deduction only the latter arguments are available. See also Dummett (1991: 202) and van Cleve (1984).

¹⁶ Of course experience may be needed to learn the concepts involved in the *a priori* claims. Notice that what is important here is not *how* we actually came to know the *a priori* claim in question, but that there is a way of justifying it that does not depend on experience. It is difficult to give a characterization of the *a priori* without offending someone. As I say below, one of the most controversial issues surrounding the *a priori* is whether it should be considered indefeasible and how this feature should be understood. For our present purposes the very sketchy characterization given above will suffice.

¹⁷ Coffa says that “one of the basic intuitions behind almost all epistemology since Plato’s is that there are two radically different types of claims: the *a priori* and the rest.” (1991: 2) The intuition can and has been challenged, but the need to take a stand on the issue has, since then, been almost unavoidable. (‘Almost’ because the epistemological relevance of the distinction can be denied. Williamson (2007: 169), for example, argues that though the distinction is clear enough, it “yields little insight.”)

When it comes to logic, the rationalists have what *seems* to be a quite direct reply to (a): *a priori* non-inferential propositional knowledge of logical principles is thought to be achieved via intuition or insight, typically delivered by a faculty of rational insight. The non-inferential character of rational insight would *appear* to make this view a good candidate for dealing with the circularity problem,¹⁸ turning the case of logic a stronghold for the rationalist account. But the well-known difficulties of giving a plausible account of what this faculty is and how it works also apply here. The empiricists, on the other hand, reject the postulation of a faculty of intuition and have long claimed that all knowledge of the world comes through experience.¹⁹ They have traditionally held that *a priori* knowledge, *if* it exists or has any application, is the result of relations between concepts, meanings, or linguistic conventions, and does not constitute knowledge of the world. The alternatives that open up here are not without challenges either: if they deny the existence or application of the *a priori*, as Mill and Quine respectively did, they have the difficult task of explaining how it is that, intuitions to the contrary, our knowledge of logical principles is, after all, empirical.²⁰ If they acknowledge the existence of the *a priori*, they face the not at all trivial task of giving a satisfactory explanation – typically by means of the notion of analyticity – of the connection between the meanings or concepts involved in logical principles and our *a priori* knowledge of them.²¹

¹⁸ Italics are meant to suggest that I don't think that rational insight views make good on their promise to deliver an inferential-free account. See chapter 5.

¹⁹ Two of the main issues of the dispute concern the source and scope of the *a priori*. This is of course a simplification of the whole debate and the positions involved, but it picks up the main ideas that have been historically influential, and it is enough for our present purposes. See Cassam (2000) for a detailed discussion of what is at issue between rationalist and empiricist theories of *a priori* knowledge.

²⁰ Quine's views are discussed in chapter 4. See Resnik (1996) and Sher (2010, 2011, 2013) for more recent defenses of related views.

²¹ For the first part of the 20th Century linguistic conventionalism was the popular way of carrying on such an explanation, until Quine came to spoil this alternative in "Truth by Convention" (Quine 1936). See section 4.1 for a brief discussion of Carnap and Quine's Carroll-inspired criticism of his view.

No doubt all these views bring in their own share of difficulties and challenges. The question is, are they hopeless? The justification problem, let alone the remaining problems, has not received all the attention one might think it merits. Lately, though, it appears to be undergoing closer scrutiny. The historically influential alternatives just mentioned, together with some new developments,²² are echoed in the ongoing debate; but it would be fair to say that most of the views being advanced are a work in progress (and presented as such). Just one point is decisively beyond dispute: the problem is very difficult. Other than that, what seems to remain mostly constant across the discussion is the need to take a stand, usually from the beginning, on the *a priori* question – not unexpectedly, of course, given the concentration on the justification problem.²³

I have come to believe, however, that it is a mistake to focus at the outset on how logical principles are to be justified. Instead, it seems to me, it is problem (c), that is, the problem about the nature of basic inferences, that should draw our initial attention. This, of course, doesn't mean that we should forget all about (a) and (b). "No treatment of any of the problems is likely to prosper which does not heed the other two" (2004b: 155), Wright warns, and indeed, the way each problem is approached is likely to have an impact on the other two. For all we know, an answer to one of them may even lead to reformulate the rest, or at least manage our expectations towards them. But one has to begin somewhere, and precisely because our treatment of one problem will affect the rest, the order in which they are approached should matter as well.

²² As for example Field's 'default reasonableness.' See Field (2000).

²³ In recent years, the question of how to account for our entitlement to the use of these inferences (our problem (b)) has received more attention, mainly furthered by Boghossian's influential work on this area.

Now, the inference problem has been largely ignored.²⁴ Getting a clearer picture of this problem will be one of the tasks of the following chapters (and indeed one that we will approach in a rather indirect way, by considering conditions that would make *impossible* for a thinker to infer). We will be mainly interested here in the epistemological aspect of the problem²⁵ and not in its psychological aspect, that is, in a study of the way subjects *actually* infer (which is also, of course, a relevant and important topic that needs to be pursued).

Among the questions about the nature of inference itself that we will be addressing, the question about the *conditions* under which a thinker is in position to perform an inferential transition will become especially important. When Matías comes to wake me up, he has come to judge that (3) I will take him to the zoo *because* he has accepted my assurance that (1) I will take him to the zoo if Sunday is a sunny day, and *because* he has become aware of the fact that (2) Sunday is a sunny day. He seems to deem his acceptance of (1) and (2) as supporting his belief in (3). It would be ridiculous, however, to expect him to know that (1) and (2) logically entail (3) – he of course won't have the concept of logical entailment. Notwithstanding this, his coming to believe that (3) is not the result of an arbitrary string of acceptances or thoughts; he establishes relations among (1), (2), and (3), and he appears to consider (1) and (2) to be his *reasons* for concluding (3). So, what is it for a thinker to perform a (deductive) inferential transition from suppositions, acceptances, or beliefs to other acceptances or beliefs? What kind of relation, if any, needs to be established, from the thinker's point of view, between the premises and the conclusion of, say, a MPP or UI inference? And what is the role, if any, of the logical principles

²⁴ Recent exceptions are, for example, Boghossian (2014), Wright (2014), Broome (2014), and Dogramaci. (2013).

²⁵ This is also the aspect stressed by the authors mentioned in the previous footnote.

when it comes to such inferences? It seems to me that we ought to begin here, with these and similar questions, before attempting a resolution of the other problems.

Why? Well, I hope to answer this question in the next chapters, but an obvious reason is that (c) seems to be a more basic problem than the problems concerning justifications and entitlements for use. To me at least, it seems clear that we need to have a better grasp of the nature of the phenomenon before attempting to explain where our entitlement to use those inferences could possibly come from.

The justification of logical principles could also use a better understanding of (c). We saw that inferences given in support of basic logical principles can be attacked on circularity grounds. But this also suggests that there is something special about these principles: they have a self-supporting character that appears to run deep. In fact, we will see that it is difficult, even for approaches that claim to be non-inferential or direct, to elude inferential uses that accord with the very principles meant to be justified. Because of this, justification attempts cannot actually turn their backs on the nature of the basic inferences themselves. Or, at any rate, it seems to me that they should not.

There are at least two other additional advantages to consider. One of them concerns the normativity of logic. We saw previously that most authors simply assume, against Harman's views to the contrary, that we perform deductive inferences and that there is a normative connection between logical principles and such inferential uses. The very formulation of the entitlement problem and the understanding that an adequate justification of basic principles could help ground the corresponding inferential uses – as opposed to uses that accord with, say, well-known fallacies – reflect these assumptions. Harman, on the other hand, is struck by what we actually seem to do when we infer or, as he prefers to say, when we reason. He is mostly

preoccupied with pointing out the differences between reasoning practices and logical inferences, which are ultimately meant to show that logic and reasoning have little or nothing to do with each other.

Now, beginning with (c) gives us some leeway to remain, at least initially, more neutral on the normativity issue. Questions regarding the nature of inferring allow us to explore the possibility of a normative connection rather than take it for granted. As the examples of the previous section suggest, we do sometimes seem to conform to logical principles when we infer, and those cases also deserve to be taken into account. But even if we forget all about the cases stressed by Harman (see pp. 6ff above) and focus only on those cases where our inferences *appear* to accord utterly with basic logical principles, a discussion of (c) need not commit us to a normative connection. We can think of such connection as a working hypothesis: it might well be that careful consideration of the nature of inferring will indicate that we cannot accommodate the normativity hypothesis, or that we don't "employ" logical principles when we infer. If that were the case, Harman's view would be strengthened: even the cases that more strongly suggest that there must be a normative connection – that is, the cases where there seems to be a perfect fit between our inferences and basic logical principles – would in the end confirm that it is merely apparent.

I suspect that at least part of the resistance to Harman's views comes from the thought that – though there may be many cases where an unsatisfactory conclusion leads to take premises back, and we are clearly disinclined to attempt the futile task of drawing all the logical consequences of our beliefs, and so on – we *do* seem, on occasion, to infer according with logical principles in our daily lives. And precisely because Harman's strong rejection of any relevant connection between logical principles and our reasoning practice is built on cases where the two come apart, the

impression survives his arguments. In any case, whether this is the case or not, (c) gives us a chance to see what is going on with the normativity issue without simply taking it for granted. That alone would be an improvement for Harman and those who find his views compelling.

The other advantage is that by setting the justification problem aside, the issue of the *a prioricity* of the logical principles need not be an immediate concern. Thinking about the case of logic as a subspecies of the more general problem of *a priori* knowledge has not been, it seems to me, particularly helpful. Emphasizing the *aprioricity* of logic has led people to focus on the traditional problems of the *a priori*, most notably whether the *a priori* should be construed as admitting revisions or not,²⁶ and whether it could be made compatible with naturalism.²⁷ I am certainly *not* suggesting here that the question of the *a prioricity* of logic should be ignored; it is an important question in its own right and deserves an answer. But I do believe that it would be better approached once we have dealt with other, more basic, questions that concern the nature of basic inferences and their relationship with the logical principles they appear to accord with. Let's not forget that the mere mention of the *a priori* tends to stir intense emotional reactions on both sides of the divide. All this talk about the *a priori* has, in my view, obscured what is really

²⁶ Euclidean geometry is usually mentioned by the critics of the *a priori* as a case that was thought to be *a priori* and turned out to be mistaken once non-Euclidean geometries were used to describe physical space. I think, however, that the issues of *aprioricity* and unrevisability are different and should not be conflated. (It has been argued that drawing this distinction is an unwise move for the *apriorist*. See, for example, Kornblith (2000: 67-70), and (Devitt 2011b: 283, fn. 14). Whether unwise or not, there seems to be here a legitimate distinction to be drawn. See Rey (1998: 27-29) for a forceful defense of the distinction and Chalmers (2011) for discussion of this issue in the context of Quine's "Two Dogmas.") On the other hand, there are reasons for thinking that we cannot draw a straightforward analogy between the case of geometry and the case of basic logic. See chapter 5 for a bit more on this.

²⁷ Devitt is an example of someone who rejects the *a priori* on the grounds that it is unexplained and unexplicable, and, in any case, not needed. He follows Quine in defending epistemological naturalism, a view that, as the title of Devitt (2011b) indicates, leaves "no place for the *a priori*." See Devitt (1998, 2005, and Devitt 2011b for a summary and expansion of his views on the issue).

distinctive about the case of logic, what sets it apart from all the other plausible candidates for the *a priori*, like mathematical and conceptual statements.²⁸

If the arguments so far have not been persuasive enough, we should also be reminded that we weren't doing so well with the justification problem after all. In my own case, what led me to see the need to set the justification and entitlement problems aside and focus on (c) is Kripke's peculiar reading of Lewis Carroll's note "What the Tortoise Said to Achilles." Though Kripke does not distinguish between the justification, the entitlement, and the inference problem, we will see that his reading of Carroll calls our attention to the conditions for inferring. Moreover, it also puts in sharp perspective what makes the case of logic especially challenging, generating problems for the most influential accounts on the justification of basic logical principles.

So maybe those who weren't moved by the previous arguments could still be persuaded in the course of the following chapters.

1.3 The Carroll-Kripke Diagnosis

What is distinctive, if anything, of inferences that accord with basic logical principles? In a series of seminars and lectures Saul Kripke has claimed that what makes the case of logic special has been brilliantly and beautifully put to us by Lewis Carroll in his classic "What the Tortoise said to Achilles." Since its publication in 1895 Carroll's three-page note has been closely scrutinized and widely discussed, and it would seem that hardly anything new could be said about it. Kripke, though, seems to have a different reading that, independently of whether it captures Carroll's original intentions, in my opinion has an impact on the justification and

²⁸ A stock example for a conceptual statement is "all bachelors are unmarried men."

entitlement problems discussed in the previous section. Since Kripke's seminars and lectures on the topic are unpublished, in chapter 2 I will go over his arguments in detail. But the main point is the following: Carroll's note has been commonly understood as showing that we cannot view the rules of inference as extra hypotheses or premises that need an independent justification; if that were the case, we would fall in an infinite regress.²⁹ To Kripke, however, it shows something about the very nature of the topic: that certain basic logical principles cannot be *adopted*. The point can be expressed in the following way: if a subject already infers in accordance with basic logical principles, no adoption is needed, and if the subject does *not* infer in accordance with them, no adoption is, as we will see, possible.

In my view, what renders Kripke's reading interesting is that by making us consider the conditions under which it would be possible for a thinker to perform, say, a MPP or UI inference (or rather, by setting conditions that would make impossible for a thinker to infer), it also forces us to consider what kind of connections could be established between the thinker's inferring process and the logical principles themselves.

1.4 Brief Outline

In chapter 2 I summarize Carroll (1895) and briefly review some of its most influential interpretations (section 2.1). The rest of the chapter is devoted to Kripke's interpretation of Carroll's note and the moral he draws from it. In section 2.2 I introduce and discuss the adoption problem, and in section 2.3 I clarify certain aspects of it.

²⁹ For an overview of Carroll's note and a review of some of its main interpretations, see chapter 2, section 2.1.

In chapter 3 I introduce basic rules of inference and argue that they cannot help us overcome the adoption problem. I also compare Kripke's interpretation of Carroll with his famous interpretation of Wittgenstein in *Wittgenstein on Rules and Private Language* (*WRPL* from now on), and discuss some points of connection between Wittgenstein's and Kripke's views on the relation between logical principles and inferring.

The following three chapters are meant to discuss, in view of the adoption problem, versions of what have perhaps been the most influential proposals for the justification of basic logical principles: empiricism *a la* Quine, rational intuition, and concept-constitution.

In chapter 4 I examine Kripke's use of the adoption problem to criticize Quine's views in "Two Dogmas of Empiricism." Since Kripke argues that Quine fails to see that the same objections he directs against Carnap (in Quine 1936 and 1954) can be applied to his own point of view in "Two Dogmas," I start with a brief review of Carnap's proposal and Quine's famous criticism of them. I continue with Quine's "Two Dogmas" views, examine Kripke's criticisms, and claim that Quine's argument against Carnap should not be identified with Kripke's argument against Quine. In the final section I discuss Kripke's Carroll-inspired metaphysical argument against the Cartesian doctrine on the creation of eternal truths and propose a related epistemological version of the argument.

In chapter 5 I critically consider Kripke's intuition-based proposal for accounting for basic logical beliefs sketched in his talks and seminars and argue that it should be rejected because it overlooks his own adoption problem. I also note some possible variations of the proposal and claim that they do not fare better.

In chapter 6 I focus on Boghossian's concept-constituting account. I discuss in detail his proposal for dealing with problems (a) and (b), and some of the criticisms it has received. I argue

that the main problem of this proposal is that it fails to accommodate the lessons that can be drawn from the adoption problem concerning the nature of inferring.

In chapter 7 I reconsider the significance of the adoption problem and focus on Ryle's famous distinction between *knowing how* and *knowing that*,³⁰ which has been under attack in recent years (mainly in the works of Stanley and Williamson).³¹ In section 7.2 I argue that "intellectualism," the claim that *knowledge how* can be defined in terms of propositional knowledge (*knowledge that*), is not an option when it comes to basic deductive inferences and that some sort of *knowing how* is needed for an account of both propositional knowledge of basic logical principles and inferential uses (Kripke's interpretation of Carroll will play a key role in my argument). I then discuss two different ways of understanding the nature of basic inferences, neither of which assigns a fundamental role to (explicit or implicit) propositional knowledge of basic principles: a view in which deductive inferential rules are 'hardwired' in us and govern our inferences without being represented, and a view where basic inferential acts are taken as a primitive ability, not internally informed by logical principles, though encouraged and improved by our direct immersion in inferential practices. I note problems for both proposals, claim that the former should be rejected, and make a tentative suggestion for making room for the normativity of logical principles in the context of the latter. Finally, I argue that independently of what it is ultimately to be made concerning the nature of basic inferences, the adoption problem already indicates the need for rethinking the justification and entitlement problems (as introduced in section 1.2.) and what we hoped to achieve with their solutions.

³⁰ Ryle (1945-46). Ryle's treatment of Carroll (1895) will be discussed chapter 7.

³¹ See Stanley and Williamson (2001) and Stanley (2005, 2011a, 2011b, 2013).

CHAPTER 2

Carroll's Regress and the Adoption Problem

Since its publication in 1895 Lewis Carroll's note "What the Tortoise Said to Achilles" has generated a great deal of attention on the part of philosophers. In Carroll's own words (or rather the Tortoise's) the colloquy between the Tortoise and Achilles provides "a lot of instruction ... for the logicians of the Nineteenth Century" (p. 280). Yet the twenty-first century finds us with not much agreement about what the instruction was supposed to be about. Any rapid overview of the current literature in the epistemology of logic will show that interest in the piece has not receded, and there is little doubt about its importance. But why exactly it is thought to be insightful is not so clear.

In what follows I will go over Carroll's note and mention some of the old and current readings of it (section 2.1). My objective here will not be exegetical, trying to come up with the 'correct' interpretation of Carroll's original intentions. Instead, I want to discuss in detail Kripke's reading of it and the moral he draws (section 2.2).³²

³² Unfortunately, Kripke's material on the nature of logic is unpublished. I have had access to the tapes and a transcription of the first two lectures of his 1974 Princeton seminar (the first, as far as I know, comprehensive presentation of the material) and to a transcription of the lecture "The Question of Logic," given at Pittsburgh in 1974. I also attended a seminar at the Graduate Center in the fall of 2006 where some of the material was presented.

Writing about unpublished material is a tricky business for both the reader and the writer. The reader may wonder whether the presentation of the original views is accurate and fair. The writer must be careful not to step in too much without saying so, since the reader will have no way of comparing her account with the original views. I have tried to explicitly say which points are solely my own and to reconstruct Kripke's main arguments as accurately as I could. However, Kripke's presentations are mainly concerned with logical revisionism (with an especial interest in Putnam's arguments in favor of

2.1 Carroll's Note

Carroll's dialogue between Achilles and the Tortoise is beautifully done and a summary of its main points will surely do it an injustice. Nevertheless, here is a brief outline (with no intention of emulating Carroll's wit and literary talents).

The Tortoise brings up the following inference for Achilles' consideration:

- (A) Things that are equal to the same are equal to each other.
- (B) The two sides of this Triangle are things that are equal to the same.
- (Z) The two sides of this Triangle are equal to each other.

and kindly asks him to write it down. She then points out that someone might accept the inference as valid and not accept A and B as true, and that someone else might accept A and B as true and not accept the hypothetical:

- (C) If A and B are true, Z must be true.

quantum logic, though he also discusses other logics), which is not the topic that occupies me here. In particular, the 'adoption problem' is usually buried among a set of worries and problems affecting the revision of classical logic (or at least certain ways of promoting such revision), and it is not connected explicitly with other aspects of the epistemology of logic. In what follows I emphasize its relevance for the problems discussed in the previous chapter.

Neither of them, according to the Tortoise, “is as yet under any logical necessity to accept Z as true.” She then asks Achilles to regard her as someone who accepts the premises but not the hypothetical, and challenges him to force her, “logically,” to accept Z as true.

Achilles appears to agree with the Tortoise’s diagnosis and takes up the challenge. The first thing he does is to ask her to accept the hypothetical C. She gladly agrees, but insists that it be written down as a third premise (“Whatever *Logic* is good enough to tell me is worth *writing down*”); after all, one may fail to see its truth. Achilles writes it down and argues that now she *must* accept Z since, he says, it follows logically: if A and B and C, Z must be true. The Tortoise points out, however, that this is *another* hypothetical that one might fail to recognize as true:

(D) If A and B and C are true, Z must be true.

Achilles admits that this may be the case and invites the Tortoise to grant D. Once more, the Tortoise agrees but wants it written down as another premise. At this point Achilles enthusiastically claims that they have reached the end of the race-course: if “you accept A, B, C, and D, *of course* you accept Z.”

In an apparently innocent way the Tortoise inquires what would happen if one still were to refuse to accept Z. Achilles reply is that, “Logic would take you by the throat, and force you to do it! [...] Logic would tell you ‘You can’t help yourself. Now that you’ve accepted A and B and C and D, you must accept Z!’ So you’ve no choice, you see.” (p. 280)

The threat, however, doesn’t seem to materialize since, as the Tortoise points out, Achilles has produced yet another hypothetical that needs to be written down ...

Months later they are seen in the same place, still writing hypotheticals down.

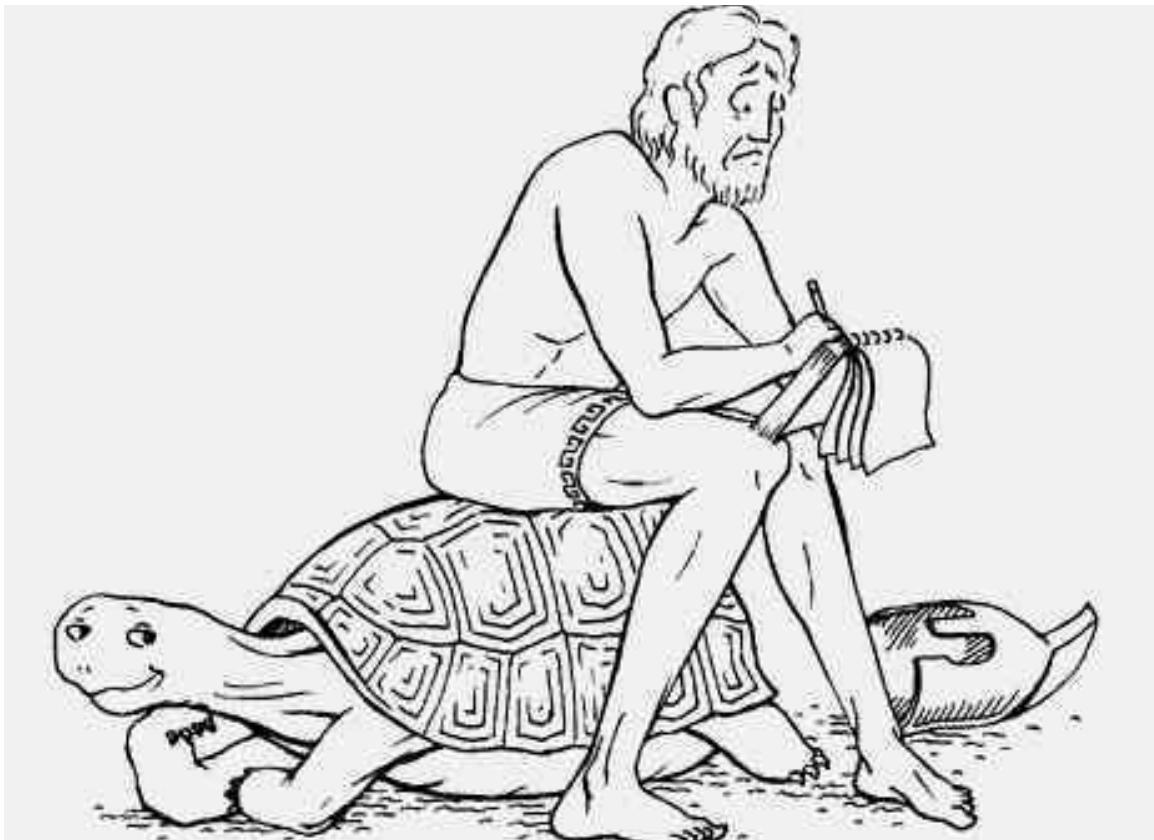


Fig. 1 Achilles and the Tortoise.

What is the point Carroll is trying to convey with this dialogue? In a letter explaining the problem to an editor of *Mind*, who was having second thoughts about the article, Carroll writes:

My paradox ... turns on the fact that, in a Hypothetical, the *truth* of the Protasis, the *truth* of the Apodosis, and the *validity of the sequence*, are 3 distinct Propositions.

For instance, if I grant

(A) All men are mortal,

and

(B) Socrates is a man,

but *not*

(C) The sequence “If all men are mortal, and if Socrates is a man, then Socrates is mortal” is valid,

then I do *not* grant

(Z) Socrates is mortal.

Again, if I grant C, but *not* A and B, I still fail to grant Z. Hence, *before* granting Z, I must grant A and B and C.

(D) If A and B and C be true, then Z is true.

Now suppose I *deny* this last sequence to be a valid one? Suppose I say “I grant A and B and C, but I do *not* grant that I am thereby *obliged* to grant Z.” Surely, my granting Z must *wait* until I have been made to see the validity of this sequence: i.e. in order to grant Z, I must grant A, B, C, and D! And so on. I think you will find that it goes on like “the house that Jack built.” (Dodgson 1977: 472)

And this explanation was persuasive enough, in view of the fact that the editor finally published the note. The letter stresses that recognizing “the validity of the sequence” plays a crucial role in the argument, but no explicit morals are drawn here either.

Probably the note’s most popular diagnosis is that both Achilles and the Tortoise neglect the distinction between a logical implication (or entailment) and a rule of inference.

A rule of inference as, for example, MPP (actually, Carroll’s inference requires quantificational logic, since it is a case of what is sometimes called “Universal Modus Ponens,”³³ but let me simplify here),

³³ One way of formalizing it would be: $(\forall x)(\forall y)(\forall z)(Exz \wedge Eyz) \rightarrow Exy$, $(\exists z)(Eaz \wedge Ebz) / \therefore Eab$. Modus Ponens (or Modus Ponendo Ponens) and Universal Modus Ponens have been classified as mixed

$$\begin{array}{c}
 p \rightarrow q \\
 \\
 p \\
 \\
 \hline
 q
 \end{array}$$

licenses the detachment of the conclusion – q above – from the premises (MPP is also called “rule of detachment”). If valid, as is the case here, the conclusion is true whenever the premises are true (under an interpretation). Rules of inference are means for obtaining new statements or propositions from previously accepted ones; without them a formal system would be unable to produce derivations. But the rules themselves are not truth bearers.

A logical implication, on the other hand,

If P then Q and P logically imply (or entail) Q

states, in the metalanguage,³⁴ a logical relationship that obtains between statements or propositions in the object language: that Q logically follows (necessarily) from P and if P then Q .³⁵

hypothetical syllogisms since antiquity. They are ‘mixed’ because they include hypothetical and categorical premises. For the history of the early development of Modus Ponens and its classification, see Bobzien (1997 and 2002).

³⁴ I’m using capital letters here to represent formulas of the object language; in the case under consideration, the propositional variables p and q . The schema of the rule of inference can be expressed in the metalanguage in the same way, using capital letters to represent formulas of the object language.

³⁵ Some formal systems, say, of modal logic, allow rules such as necessitation, that is, from A to infer $\Box A$. Such a rule is logically valid in the sense that, if the premise(s) are theorem(s) of the system, so is the conclusion. Usually such a rule need not be postulated and can be obtained as a derived rule – as a metatheorem about what is derivable in a system. If such a rule is allowed, it could not be used for deducing conclusions from premises that may not be logically valid. Church (1974: 148) – who appears to have trouble in one of his systems obtaining necessitation as a derived rule – regards it as objectionable to

While these notions are intimately connected, because a rule of inference is valid if and only if the premises imply the conclusion (I'm not making judgments here as to which notion is more basic), the notions are distinct.

With this in view, consider what Achilles and the Tortoise are doing. Instead of drawing the conclusion on the basis of the premises and according to the **MPP** rule, the Tortoise turns the relationship between premises and conclusion into a logical implication, or, as Carroll's calls it, a hypothetical:

(C) If A and B are true, Z must be true,

which could be expressed as:

A, B logically imply Z.

[Carroll's hypothetical could be expressed in the object language using a strict conditional or modal operators. Note also that the **MPP** rule can also be turned into a corresponding material conditional, that is, $((p \rightarrow q) \ \& \ p) \rightarrow q$ (in many systems it can be shown by the deduction theorem that if $\Gamma, A \vdash B$, i.e. if B is derivable from Γ and A, then $\vdash A \rightarrow B$ is derivable from Γ alone). For the rule to be valid the conditional has to be a logical truth (and vice versa), true under any assignment of truth-values to its non-logical terms. Merely adding the corresponding material conditional as an extra premise would be weaker than adding the hypothetical or logical implication, unless it is explicitly stated that the conditional is a logical truth. But note as well that

take necessitation as a primitive rule. He claims that only rules that derive conclusions that follow from the premises should be allowed in a proper system. (My thanks to Saul Kripke for pointing this out to me.)

even if this is *not* stated and the corresponding material conditionals of the Tortoise's potential inferences are simply added as premises at each stage, the regress would still arise, with no talk of validity, logical implication, or logical truth involved, since there won't be a procedure that licenses the detachment of q . Raising the regress might be unmotivated, but that is a different issue.]

The Tortoise brings up the hypothetical as an extra element that needs to be granted before the conclusion is drawn. But, since we can fail to note its truth, she refuses to draw conclusion (Z) and opts to “wait” until she has been “made to see” the truth of the hypothetical. Both the Tortoise and Achilles seem to agree that this is a good reason for its inclusion as an extra premise. And indeed, once we have agreed that not only the premises but also the hypothetical need to be granted for the conclusion to be drawn, adding it as an extra premise appears to make sense (the hypothetical would be, on this view, a hidden premise or assumption of the inference, made explicit by its introduction as a premise). The process then repeats itself at each stage and no conclusion is ever drawn.

One way to see what is going on in this exchange is to point out that by writing down (C) (and (D), (E), and so on) on a par with (A) and (B), Achilles and the Tortoise neglect the distinction between rules of inference and logical implications. A rule of inference calls for the detachment of the conclusion, and by turning the rule of inference into an implication – which only states the relation between the premises and the conclusion – they are defeating its whole purpose. According to this interpretation, it is precisely their neglecting this distinction what causes the regress (they are mixing levels of language as well, though this could be avoided by introducing

the corresponding conditional instead).³⁶ This reading – let’s call it the *rule-as-premise interpretation* – is often seen as a solution as well: we stop the regress once we note that turning rules into extra premises of the inferences they inform violates the distinction.³⁷

Some authors, while acknowledging the problem that the *rule-as-premise* interpretation generates (or even admitting that it may be what Carroll had in mind), see it as signaling other problems. Hanna, for example, stresses this interpretation, though he thinks that a “deeper interpretation” in which what is questioned is “what grounds valid deduction *itself*” is also possible (Hanna 2006b: 57ff). Ryle states that the “principles of inference are not extra premises” (1945-46: 7) and then goes on to argue that the problem lies in the obliteration of the distinction between knowledge-that and knowledge-how: Achilles feeds the Tortoise with the former, but only the latter is relevant in such a situation.³⁸

³⁶ Wisdom (1974) thinks that the problem is generated by the need to involve different orders of language: someone trying to justify an inference is in danger of falling into a regress because such justification would have to proceed on a second order level, and the justification of the justification on a third order level, and so on. Each level can only be justified from a higher order, and so someone seeking an ultimate justification would go on forever.

Hanna (2006b: 56) and others also point out that Achilles and the Tortoise are mixing levels of language.

³⁷ See for example Rees (1951), Geach (1965), Thomson (1960), and Smiley (1995). Thomson mentions the failure to make this distinction, but goes on to argue that the regress “is just an infinitely long red hearing” (p. 105): since the Tortoise has accepted A and B, she is ‘under a logical necessity’ to accept Z (p. 96).

Russell (1903: §38) also makes the point that the notion of ‘therefore’ (as in ‘p, therefore q’) should be distinguished from the notion of ‘implies’ (as in ‘p implies q’). He claims that while the former is a relation between two asserted propositions, in the latter neither ‘p’ nor ‘q’ is asserted but merely ‘considered.’ Drawing this distinction is supposed to be the “first step in answering Lewis Carroll’s puzzle” since the Tortoise treats both cases as if they were the same. Whatever the merits of this proposal as an answer to the puzzle, there seem to be independent reasons for rejecting it (reasons noted by Russell himself), since one of its consequences is that a proposition would be changed by its assertion. Then in the inference ‘p, if p then q, therefore q’, ‘p’ seems to mean two different things in the first premise and the antecedent of the conditional. That by itself gives reason enough for suspecting that something has gone wrong.

³⁸ Priest (1979: 291 and fn. 4) also argues that the main moral of the note is that implication (entailment) and inference should be distinguished. He stresses that while an implication is a relation between sentences (the premise(s) and the conclusion), drawing inferences is an action. Thus, Carroll’s note “shows that no extra entailments, taken as premises, can force a person to infer i.e. to do something.” He credits Ryle (1950) for drawing this moral.

Different aspects or applications of the note are underscored by other interpretations. Blackburn (1995) and Schueler (1995), for example, and following them Engel (2005), argue that the point the note raises is whether logic can move the mind. They draw an analogy between theoretical and practical reasoning³⁹ and try to make a case for the possibility of logical *akrasia*. The phenomenon of *akrasia* or weakness of will is that of an agent choosing to do what she judges she ought not to do, and it is ordinarily thought to arise in the domain of action. Analogously, logical *akrasia* is the phenomenon of a subject recognizing the truth of a logical law, or the validity of an inference, but failing to infer accordingly. (See Engel 2005: 27)⁴⁰

Another possibility is to claim that, contrary to what she says, the Tortoise has not really accepted premises A and B. If her acceptance were indeed sincere, her refusal to accept the conclusion would alternatively indicate that there is at least one premise that she does not understand. It is simply not possible to understand and truthfully accept A and B, and reject Z. So she is either lying or confused.⁴¹

Boghossian, on the other hand, emphasizes that the Carrollian argument “is meant to raise a problem for the *justification* of our rules of inference – how can we justify our belief that Modus Ponens, for example, is a good rule of inference?” (Boghossian 2008a: 128. As we will see, he draws other morals from the note as well.)

And of course we should not forget the most influential application of the note, Quine’s “Truth by Convention” (Quine 1936: 103-104; and also Quine 1954), where he famously argues that Carroll’s regress shows that logical truths cannot be given by means of conventions.

³⁹ Such an analogy is already in Ryle (1945-46). See chapter 7 for discussion.

⁴⁰ The main moral Blackburn draws from Carroll’s note is “That there is always something else, something that is not under the control of fact and reason, which has to be given as a brute extra, if deliberation is ever to end by determining the will.” (1995: 695)

⁴¹ See for example Black (1970).

In the following chapters we will come back to some of these morals and applications. Let's now move on to Kripke's reading.

2.2 The Adoption Problem

As I have already suggested in the introduction, the main moral Kripke draws from Carroll's note is that certain basic logical principles cannot be *adopted*.⁴² But before getting there let me reconstruct the main idea in the following way.

Assume we are trying to get someone to adopt a very basic logical principle, like UI or MPP. The subject, let's call him Harry, has no notion of the principles in question and has *never* inferred in accordance with them.

Imagine now that we think it would surely be a good idea for Harry to *adopt* them. By 'adopt' here we mean that the subject, Harry in this case, picks up a way of inferring according to, say, UI, something he wasn't able to do before, *on the basis* of the *acceptance* of the corresponding logical principle. As we go on we will need to be more precise, but for now let's stick to this rough characterization to get a flavor of the problem.⁴³ (Keep in mind that 'adoption' as it will be

⁴² The term 'adoption' was already part of the literature Kripke was discussing (in particular, Putnam (1962 and 1968)). He complains that phrases such as 'adopting quantum logic' or 'the adoption of intuitionistic logic' were commonly used as if the term 'adoption' had a well-known and clear sense. Putnam, for example, says: "Suppose we are willing to adopt the heroic course of changing our logic. What then?" (1968: 179) and "... the *adoption* of intuitionistic logic as opposed to 'classical' logic amounts to systematically forswearing certain classically valid inferences" (1962: 51. My emphasis.) Kripke repeatedly claims that he doesn't know what 'adoption' means in this context and doesn't attempt to give a precise formulation.

⁴³ This is my first approximation to the notion of adoption; for further clarification, see section 2.2.2. Instead of a logical principle we could give Harry a rule of inference, or the principle or rule plus some relevant examples. These cases are considered in chapter 3.

We could also claim that a subject *S* has 'adopted,' say, MPP if, on the basis of the acceptance of the MPP implication or rule of inference, *S* is led to detach or assent to the relevant conclusion when she assents to the premises of a MPP instance, something *S* didn't do before.

understood here does not simply consist in picking up a basic inferential practice, but doing so by means of the acceptance of a logical principle.)

Someone like Harry – that is, someone who has never inferred in accordance with the UI principle – is, admittedly, not someone we would likely encounter around the corner. It should be clear from the beginning that I am *not* suggesting that thinkers start out or find themselves at some point in a situation that calls for their *adoption* (in the sense given above) of basic principles like UI. The proposal is, instead, to think about Harry’s case more as a thought experiment than a real scenario.⁴⁴

So, what can be said to help Harry out? Suppose we say to him: “Look, Harry, you need to accept the following: ‘Universal statements imply each of their instances;’ or alternatively, ‘the predication of a property (say we ignore relations) to all members in a domain entails that any particular member has that property.’ What I mean by this is simply that if some property applies to *all* the things we are considering, it logically follows that the property applies to *each particular* thing in the group. Don’t worry, this principle is true and I guarantee that you will find it useful.”⁴⁵

⁴⁴ It may be objected that the example is too fantastic and far-fetched to be of interest. I concede to the strangeness of the case, but is still instructive. We are seeking an understanding of the nature of inferring, and discussing extreme circumstances can give us a better grip on its conditions (after all, Carroll’s exchange is not very plausible either!). As I said, Harry’s example is not meant to suggest that we “adopt” basic principles like UI (that is, that we start inferring according to UI just by being told the UI principle), but to work as a *reductio*: inferring according to UI would be impossible *merely* on the basis of the acceptance of the UI principle. See also fn. 48 below and section 2.2.2.

⁴⁵ Nozick (1981: 227ff) argues that knowledge is not closed under universal instantiation. He discusses specific cases where, according to him, a UI premise is known, but an instance is not, even where someone sees the validity of the inference and draws the conclusion. However, the point is not that the logical implication doesn’t hold, nor that someone who believed the premise would not believe the conclusion, but rather that (supposedly) known logical implication does not always preserve knowledge. So the conclusion, even though properly derived by UI, need not be knowledge even if the premise is known. Nozick’s views on closure in general and in this case in particular are controversial and I won’t discuss them here. My thanks to Saul Kripke for raising this issue to me.

“Fine,” says Harry, “I accept your suggestion, you seem to know a lot about this stuff, and I will take your word for it.”

Suppose we now want to make sure he has truly adopted UI and run the following test. We ask him to read a review of *Madagascar* that, among other things, says that, “all the animals in the movie talk,” and describes Alex the lion as “the most likable animal featured in the movie.” We then ask Harry whether Alex the lion talks.⁴⁶ Much to our dismay, he replies: “I have no idea. I told you I haven’t seen the movie.” We insist: “haven’t you noticed that the review says that all the animals in the movie talk?” “Yes, I have.” So we inquire, “do you doubt the review’s claim?” “No,” he says, “but you didn’t ask *that*; you asked me if the lion Alex talks.”

By now we are pretty sure he has adopted nothing and say: “come on Harry, we talked about this, didn’t you say that you would accept that ‘all universal statements imply each of their instances?’” Perhaps, we think, we didn’t give him enough information, we haven’t told him how a universal statement looks, and maybe he just failed to recognize it as such. So we say: “Maybe you haven’t noticed, but ‘all the animals in the movie *Madagascar* talk’ is a universal statement.

⁴⁶ Universal instantiation is usually understood as ranging over a domain given by the subject term (in this case, the animals in the movie *Madagascar*), and this is the way it will be understood here. The domain may not always be spelled out, but it is implicit in the context of most ordinary assertions. Lately, there has been a lot of discussion about whether quantifiers could range over absolutely everything there is (see the introduction to Rayo and Uzquiano (2006) for an overview of arguments in favor and against this view). Even if completely unrestricted quantification were possible, defenders of the view would agree that this is not always the way ordinary speakers commonly understand it.

Universal instantiation fails in free logics because constants (or other singular terms) can denote non-existent objects. So, even when every member of the quantifier’s domain satisfies the predicate, a constant may not denote a member of the domain. One of the motivations for free logic comes from the need to account for fictional discourse. Maybe cases involving fictional or mythical characters, could be handled by appealing to an ontology of fictional or mythical characters, as is done in Kripke (2011b and 2013). This could be a problem for the example, since Alex and the animals in *Madagascar* are fictional characters. I will assume, however, that the domain is not empty and won’t worry about empty names (those that find this assumption problematic can think of *Madagascar* as a documentary film!).

In general, I want to set aside doubts about the general applicability of the principle. I don’t think anything hangs on this: we could still ask if it would be possible to get someone to perform a universal instantiation for those cases that we consider beyond doubt.

They often have the form ‘All x ’s are P ’s,’ where the x ’s are the animals in *Madagascar* and P is what we are saying of them, the predicate that we are applying to them, in this case ‘talk’.” “So,” we go on, “as you have agreed, *all* universal statements imply each of their instances, and *this* particular universal statement, “all the animals in the movie *Madagascar* talk,” implies *this* particular instance, “Alex the lion talks,” since a lion is obviously an animal.”⁴⁷

“Does it?” says Harry, “I haven’t heard of a lion that is not an animal, and I believe you when you say that all universal statements imply all of their instances, but I don’t see the last thing, that I should be able to tell whether the lion Alex talks. Why is that? That review didn’t say anything about the lion talking. Maybe I should just go and see the movie.”⁴⁸

What can we say to that? It is clear that the problem could not just be that lack of information prevented Harry from recognizing a universal statement. The difficulty is that to be able to recognize it *as* a particular universal statement he has to perform a universal instantiation, something that, by hypothesis, he doesn’t do. Nothing but more trouble has been gained with this exchange: he now has to pick “all the animals in the movie *Madagascar* talk” as an instance of a universal statement, then infer from “All universal statements logically imply each instance” that “all the animals in the movie talk implies each instance,” and then, assuming he read that Alex the lion is one of the animals in the movie (otherwise, if it just said that Alex is a lion, we would

⁴⁷ The inference we are trying to get Harry to perform is, as in Carroll’s note, a case of Universal Modus Ponens (and a mixed hypothetical syllogism), though this case is simpler since it doesn’t require relations and existential statements. In symbols the inference would be $(\forall x)(Ax \rightarrow Tx), Aa / \therefore Ta$.

⁴⁸ Would it be possible for someone like Harry to understand predication at all? Maybe not. I don’t have a settled view about this: a possible reaction is that we can think of him as someone that simply operates at a propositional level, but would Harry even be able to understand ‘Alex is a lion’? It isn’t clear to me. In the *Tractatus* Wittgenstein thought that in a proper language the role of the quantifiers should be occupied by disjunctions and conjunctions. Whatever may be the case, the point remains that our telling him the UI principle would not help him infer or understand such assertions. My thanks to Alberto Moretti and Luis Robledo for emphasizing this to me.

need another inference here), recognize it as an instance of the universal statement and infer that “Alex the lion talks.”⁴⁹

2.2.1 Harry and the Tortoise

Kripke thinks that the main moral to be drawn from Carroll’s note is that certain basic logical principles have a special status: unlike other statements and principles, they cannot be adopted.⁵⁰

⁴⁹ In his lectures and seminars, Kripke argues that before considering whether we can adopt an alternative logic we should see if we could adopt something more familiar. He says:

Let’s try to think of someone – and let’s forget any questions about whether he can really understand the concept of “all” and so on – who somehow just doesn’t see that from a universal statement each instance follows. *But* he is quite willing to accept my authority on these issues – at least, to try out or adopt or use provisionally any hypotheses that I give him. So I say to him, ‘Consider the hypothesis that from each universal statement, each instance follows.’ Now, previously to being told this, he believed it when I said that all ravens are black because I told him that too. But he was unable to infer that *this* raven, which is locked in a dark room, and he can’t see it, is therefore black. And in fact, he doesn’t see that that follows, or he doesn’t see that that is actually true. So I say to him, ‘Oh, you don’t see that? Well, let me tell you, from every universal statement each instance follows.’ He will say, ‘Okay, yes. I believe you.’ Now I say to him, “‘All ravens are black’ is a universal statement, and ‘This raven is black’ is an instance. Yes?’ ‘Yes,’ he agrees. So I say, ‘Since all universal statements imply their instances, this particular universal statement, that all ravens are black, implies this particular instance.’ He responds: ‘Well, Hmm, I’m *not entirely sure*. I don’t really think that I’ve got to accept *that*.’” (Pittsburgh lecture)

The example is then connected with Carroll’s note and both cases are treated as the same. Ryle (1945-46) discusses a related example: a student who understands the premises and the conclusion of an inference yet doesn’t see that the conclusion follows (but there are important differences between the examples as given. See chapter 7).

⁵⁰ Kripke reports that his first thoughts on the nature of logic were developed in reaction to some religious arguments (while attending a summer camp). The basic thought was this: it is sometimes argued that religion is not alone in having to accept certain principles without empirical or rational proof. Science, for example, has to assume the existence of the external world and the regularity of nature. The analogy is meant to show that all disciplines, even the scientific ones, have their own set of unproved assumptions, and so religion should not be considered peculiar in this respect. In reaction to this kind of argument Kripke reportedly argued that this cannot apply to *all* disciplines. In particular, it would not apply to logic, since basic logical principles cannot just be accepted as assumptions because as such they would be useless.

In his seminars Kripke also draws an analogy between the strategy adopted by some authors who question the legitimacy of basic logical principles (he mainly has in mind Putnam 1968) and certain

For certain basic logical principles, if we do not already infer in accordance with them, accepting or assuming them would lead us nowhere.

In particular, UI cannot be adopted (in the sense of “adoption” given before). As we have seen, Harry, who has never performed UI inferences, would not benefit from our telling him that “all universal statements imply each of their instances” because he would need to apply UI to the UI principle to get it moving.

The same goes for MPP. If someone who never inferred in accordance with MPP were to be told that “For any A and B, if A then B, and A, then B,” the subject wouldn’t be in a better position to perform a MPP inference. For the principle to be of use with any particular inference, she will need to infer in accordance with the MPP pattern that she does not use in the first place: in any particular case, she will only get to B from her premises by performing a MPP inference on the instantiation of ‘For all A and B, if A, and if A then B, then B,’ but that is exactly what she couldn’t do to begin with.^{51 52}

religious arguments: both attack their opponents by saying that though their arguments are correct as stated, they assume ‘classical logic’ or ‘reason.’ They then reject those assumptions and go on to claim that their opponents’ conclusions have been blocked. He calls this strategy “methodological ascent.” (For more on this see chapter 5.)

⁵¹ Alleged counterexamples to MPP have been given by Adams (1975), McGee (1985), and Lycan (1993). The most famous case, proposed by McGee, is as follows: “Opinion polls taken just before the 1980 election showed the Republican Ronald Reagan decisively ahead of the Democrat Jimmy Carter, with the other Republican in the race, John Anderson, a distant third. Those apprised of the poll results believed, with good reason: If a Republican wins the election, then if it’s not Reagan who wins it will be Anderson. A Republican will win the race. Yet they did not have reason to believe: If it’s not Reagan who wins, it will be Anderson.” (McGee 1985: 462). This case, however, is controversial and some people have argued that it doesn’t constitute an instance of MPP. Whatever it may be going on here, it does look much more complicated than the cases I have been considering, and I will continue avoiding problematic cases like this in what follows. I won’t discuss cases in which the antecedent of the material conditional is not relevant to the consequent either. Questions of relevance have motivated the advent of relevance logic (in fact, a family of relevant or relevance logics), which rejects the disjunctive syllogism, $p \vee q, \neg p \therefore q$ (also rejected by dialethetists as a way of dealing with the Liar paradox), and hence the material conditional is not closed under MPP. But, as I said, I won’t worry about this problem here.

⁵² Kripke mentions two other principles as being affected by the adoption problem, adjunction and non-contradiction (the list, however, wasn’t intended to be exhaustive). According to adjunction if ‘A’ is true

Now, Kripke treats Carroll's Tortoise and cases like Harry's as if they were the same. But there are important differences between the two cases.

Notice, in the first place, that while Harry is someone who has *never* inferred in accordance with UI, the Tortoise is in a different situation. She is proposing a challenge in which Achilles is supposed to "logically" force her to accept a particular conclusion (Z). As far as we know, the Tortoise has made similar inferences in the past and is aware that (Z) follows from (A) and (B), but as a part of the challenge she chooses to neutralize her own judgment and see if she can be *forced* to accept (Z). In each case she recognizes the relevant hypothetical and is willing to accept its truth as soon as Achilles asks her to do so. But in the game that she has set out to play she has to remain, let's say, impartial: nothing should be granted without being added as an extra premise. In other words, she is resisting the *act* of drawing the conclusion because she wants to be sure that every step has an appropriate justification, and this self-imposed impartiality leads her to an infinite regress of premise adding.⁵³

Harry, on the other hand, never gets to that point; he never falls in a regress. He is told what the principle is, but is unable to do anything with it from the start. He is not trying to resist the inferential act, to remain impartial until every step is justified, or play the skeptic's part. He has *never* inferred in accordance with UI and is just clueless as to what to do with the principle once it

and 'B' is true, then 'A and B' is true. The problem is that if we accept that 'A' is true and accept that 'B' is true, then adjunction is needed to conclude that one has accepted that "A is true and B is true." In the case of non-contradiction, if we try to get someone who doesn't reason in accordance with it to adopt it, the problem is that the subject could hold the law of non-contradiction itself to be both true and false (this case is similar to the 'perverse instantiator' example, see below). Michael Levin has suggested to me that simplification is also subject to this problem. In what follows I will only talk about UI and MPP. The former is Kripke's preferred example and the latter is Carroll's own example (and also the case that is most widely discussed in the literature), but of course it would be interesting to see which other principles are directly affected by this argument. That task, however, falls outside the scope of what I set out to do here. In any case, I think that the argument would be interesting even if it only worked with MPP and UI.

⁵³ It is the whole point of Carroll's piece to generate a regress, as the names of the characters participating in the exchange clearly indicate. As in the case of Zeno's paradox, there seems to be a problem here involving the possibility of movement.

is given to him. This is why logical *akrasia* could not be imputed to him; he doesn't *choose* not to infer, he simply doesn't have an alternative.

Another difference is that while the Tortoise wants to be forced to perform a *particular* inference, we are trying to get Harry to adopt a *general* principle that would apply to all inferences of the same pattern. Achilles doesn't propose a general principle to the Tortoise. In each case he gives her a particular hypothetical (e.g. (C) if A and B are true, Z must be true), which states the logical implication between the premises and the conclusion of the particular inference he is trying to get her to perform. He hopes that once the Tortoise has accepted the hypothetical in question, whether it is (C), (D), (E), etc, she will also have to accept conclusion (Z), but of course the plan doesn't work very well because she keeps adding the hypotheticals as premises.

When it comes to Harry, we are more ambitious: even though we could tell him "from 'all the animals in the movie *Madagascar* talk,' 'Alex the lion talks' follows" and he would probably take our word for it, we do not want to do that since it will only help him with *this* inference. We would simply be inferring *for* him. What we want is for him to take up an inference pattern that accords with the principle, but the UI cases are infinite and it cannot be done case by case. This is why we give him a general principle; it is supposed to *guide* his future applications.

One of the advantages of setting the problem in terms of Harry's case is that it uncovers a deeper problem than the need to distinguish between rules of inference and implications.⁵⁴ We saw that the rule-as-premise interpretation is sometimes accompanied by the idea that once the distinction is in place, the problem dissolves. But Harry's case suggests that there is also a problem with cases where no rule is added as an extra premise. In contrast to the Tortoise, Harry's problem is that he would *already* need to infer in terms of UI in order to be able to *use* the

⁵⁴ I am not suggesting here that this moral and analysis of the problem is incorrect, but that the rule-as-premise interpretation, though correct as stated, does not capture the whole of the matter.

explicit statement we are giving him. One might think that telling him the principle is, given Harry's situation, another way of actually adding an extra premise to the UI inferences he may encounter. And that seems to be what Kripke had in mind, since he says: "If he was not able to make the simple inference, "All ravens are black, therefore, John is black" (where John is a particular raven), then giving him some 'super premise' like "Every universal statement implies each instance" as another premise won't help him either." (Pittsburgh Lecture, see fn. 49) But that presupposes that he would be able to identify the inferences for which the principle is relevant, and considering that this requires UI, we can't expect him to get to that point: he just won't know *where* to add the principle as a premise (though, of course, even if we tell him, it wouldn't help). Moreover, we are not telling him the UI principle because we think that it needs to be granted in order to deduce the conclusion, but because we mean for the principle to serve as a guide in his inferential practice. And, more importantly, as we will see in chapter 3, telling him a rule of inference won't help him adopt UI either.

A further point worth mentioning is that though Carroll's inference and the inference we are trying to get Harry to perform can both be seen as cases of "Universal Modus Ponens" (at least in the formalizations proposed), the two examples emphasize different aspects: the Tortoise's troubles revolve around MPP and its role as a rule of detachment, and Harry's problem is with UI (though we just saw that a similar problem can be generated for MPP).

Kripke said in passing that UI is a nicer example, though he didn't say why and is now not sure about what he had in mind. Perhaps UI could be thought to be more basic, since MPP as a generally stated principle needs the inference pattern of universal instantiation to be of use: without it we wouldn't be able to conclude that a particular instance is a case of MPP. But this doesn't seem right, since UI also seems to require MPP. In the *Madagascar* example the domain of

the quantifier is explicit, the animals in the movie, but even when the domain is implicit, we are also presupposing a MPP inference, that is, that if something is a member of the domain, then it satisfies the relevant property or relation, and only if the antecedent holds, we infer that an individual or object satisfies the property or relation in question.

Moreover, note that what is important in Carroll's example is not that the inference itself is a case of MPP, but the intimate connection between the validity of an inference and the logical implication between its premises and conclusion (or, as I said before, the inference's associated conditional). In *The Problems of Philosophy* Russell says that the principle, "Suppose it known that if this is true, then that is true. Suppose it also known that this is true, then it follows that that is true.' ... is really involved – at least, concrete instances of it are involved – in all demonstrations. Whenever one thing which we believe is used to prove something else, which we consequently believe, this principle is relevant." (Russell 1912: 71-72) As Russell presents it, the principle itself is a MPP inference, but it can also be formulated as an implication (or hypothetical) or an associated conditional. And if we happen to believe, as Achilles and the Tortoise seem to do, that something like this principle needs to be granted for an inference to go through, we are already on the way to Carroll's regress. It does not matter whether the original inference is a MPP inference or not, what is relevant here is that MPP can be used to think *about* particular inferences in the way suggested by Russell.

For our purposes the most important difference is that the Tortoise's insistence that no step should be taken for granted places the emphasis on having an adequate *justification* for each step; and, in contrast, in Harry's case the concern is whether it would be possible for him to make the inference *on the basis* of the information we are giving him, placing the emphasis on the inference itself and the conditions under which a subject is in position to perform it.

All these differences may suggest that we are dealing with a different problem and that Harry's problem is not the one Carroll had in mind. I have already said that I won't try to exegetically determine here what Carroll's problem exactly amounts to, or what he believed its main moral to be. If Harry's case, though inspired by Carroll's note, poses a different but interesting problem, it should not be ignored.

Having said that, I should add that there is an important point that connects both cases and allows us to see them as variations of the same theme. Neither the Tortoise nor Harry actually *perform* the inference. Both are unable to use the relevant principle or hypothetical as a reason for inferring. Both claim to accept the principles or hypotheticals given to them but, either because the desire to be impartial and justified leads to the suspension of all inferring or because, since there is no inferring to begin with, they cannot draw the conclusion. In both cases the problem seems to be that if the premises of the inference weren't sufficient for them to draw the conclusion, telling them or making them accept the relevant principle or hypothetical would be of no help.

2.2.2 Some Clarifications

Let's express the fundamental point in the following way:

Carroll-Kripke Adoption Problem (AP): certain basic logical principles cannot be *adopted* because, if a subject already infers in accordance with them, no adoption is

needed, and if the subject does *not* infer in accordance with them, no adoption is possible.⁵⁵

At the beginning of the discussion we said that, “by ‘adopt’ here we mean that the subject, Harry in this case, picks up a way of inferring according to, say, UI, something he wasn’t able to do before, *on the basis* of the *acceptance* of the corresponding logical principle.” (p. 38)

What does “on the basis” mean here? Let me first say that it does *not* mean a mere causal connection, where being told the principle is simply supposed to somehow *cause* the desired behavior. Who knows, after all, what our telling Harry the relevant principle might cause him to do?

What we are looking for is for the principle to *guide* Harry’s present and future inferential practices. The principle should provide Harry with the reason for performing particular inferences: it should be *in virtue* of his acceptance of the principle that he infers in accordance with UI. A mere causal connection would be insufficient because the relation between Harry’s acceptance of the principle and his inferring in accordance with it could be fortuitous. For now, we want to exclude this possibility. Again, if the principle is going to function as a real guide, Harry’s acceptance of it should constitute his grounds for inferring according to it.

What we seek by means of the adoption of UI is then a constitutive relation between the principle and the inferential act, not just an extrinsic one. The problem is that for the principle to have this grounding role, it seems that there is no alternative than to make the relation an inferential one. How else could Harry’s acceptance of the principle guide him? He would have to deduce from it its applications: from “all universal statements imply each of their instances” and

⁵⁵ The name and expression of the problem are my own. Kripke doesn’t give any particular name to it or attempt to give a more precise formulation.

a particular universal statement, he will have to infer that a particular instance of the universal statement is implied. And, of course, the deduction involved is none other than the one he couldn't do to begin with (plus MPP).⁵⁶

A related question was asked by Hempel in the discussion section of the Princeton Seminar. He asked whether someone could adopt a non-standard logic by going through a process of psychological conditioning (Skinnerian conditioning, as Kripke recalls it) that would make him behaviorally *conform* to the alternative logic. Unfortunately, there is no record of this exchange, but Kripke recalls arguing that this would not really be an 'adoption' of an alternative logic and that we wouldn't say of someone whose inferences conform to known fallacies that he has adopted an alternative logic. The main target of Kripke's criticisms in the seminar are, on the other hand, proposals where the adoption of an alternative logic is seen as a rational process akin to the adoption of a scientific hypothesis (see chapter 4). And this is not what is going on in the case proposed by Hempel.

Given the way we are understanding "adoption" here, it is clear that such conditioning would not count as adopting UI: Harry's acceptance of the principle would not be *his reason* for inferring

⁵⁶ We assumed from the beginning that MPP and UI are basic, not derived from other principles, but it could be objected that even if a direct adoption of UI and MPP is not possible, there could be a roundabout way of getting a subject to adopt these principles. It is not hard to see, however, that we won't go very far with this. Suppose we try to get Harry to adopt the following more complex principle: $\neg(\text{Ex})\neg\text{Fx} \rightarrow \text{Fa}$, 'if there doesn't exist something for which a predicate F does not hold, then F holds for any particular instance.' Would it matter? It is clear that any other principle or rule we might suggest to Harry will have to be general and so their application will presuppose UI.

What about MPP? We could express it as $((\neg p \vee q) \wedge p) \rightarrow q$. Again, its application will involve use of the MPP inference pattern. Consider: how would Harry get to the conclusion q? Suppose he accepts $\sim(p \& \sim q)$, p, and the principle $[\sim(p \& \sim q) \& p] \rightarrow q$. From all this q follows, but a MPP inference is involved. As Russell points out and we have emphasized already, any mediated attempt to detach a conclusion will presuppose inferring in accordance with MPP.

Maybe this is what Kripke has in mind when he says that sometimes the more basic principles are "built into" the other ones in a disguised way (he mentions this in passing and says he doesn't want to get into this issue).

in accordance with it, in fact the principle itself may be playing no role. (Harry would be merely *conforming* to the principle in this case.) Still, the question remains: why construe “adoption” in this restrictive way? Why exclude cases where the acceptance of the principle plays no role?

The question touches directly on the relevance of the adoption problem. Though we may find Harry’s case improbable, the idea that it is *because* we have implicitly accepted, grasped, or internalized UI or MPP principles (or rules) that we are able to produce indefinitely many UI and MPP inferences we never considered before, has considerable intuitive appeal. Most of us, on the other hand, find it natural to think of such inferential acts as being voluntary and under our rational control (even when we may not be able to give an explicit formulation of the rule we are according to). We don’t think of ourselves as being drilled by conditioning processes or sub-personal mechanisms. But then, how else are we going to account for our capacity to perform UI and MPP inferences? We obviously did not contemplate their infinite applications. Presuming that we have somehow accepted general principles (or rules) suggests itself as the obvious way of accounting for this capacity. Or so it seems.

Of course, we were never in Harry’s situation, we did not “adopt” UI or MPP inference patterns by being told the relevant principles (or rules). Nevertheless, how we came to grasp the principles is secondary. Perhaps exposure to UI inferences in different situations caused us to implicitly accept or grasp the UI principle, perhaps the principle itself was inborn in us. In the view under consideration what’s important is that it is *because* we have grasped or accepted the principle (or rule), *because* we have some kind of mental state encoding the principle, that we are in position to perform UI inferences in a potentially infinite number of cases. This internalization of the principle, this state of acceptance (which may well be implicit) which encodes the principle

or rule is then supposed to inform our inferential behavior. And it therefore has a guiding role: it guides our inferential uses, which are in turn explained and evaluated by reference to it.

Can a view along these lines be correct? It seems to me that the justification problem became *the* problem in the epistemology of logic precisely because something like the view outlined above has been presupposed. At first blush it makes perfect sense: if we rely on such principles to obtain a large number of beliefs, we better find a way of justifying them. But the peculiar character of the difficulties facing the justification problem should perhaps have forced the question of whether the presupposed view was indeed adequate more than it has.

At any rate, Harry's case and the adoption problem are meant to challenge views on the nature of inferential acts where acceptance plays a fundamental role. Other views, such as the one considered by Hempel, where intentional states are not the basis of our inferential acts, won't be touched by Harry's troubles. And for this reason they will be considered last.

An important distinction that has been implicit in our discussion so far is the distinction between *logica utens* and *logica docens*.⁵⁷ As the Latin suggests, the distinction was already known to medieval writers, but was reintroduced in the literature by Peirce.⁵⁸ *Logica utens* (logic in use or action) is the logic that a subject or community of subjects use when performing inferences, the "logic in possession," and thus concerns the *practice* of logic. In contrast, *logica docens* (theoretical or pedagogical logic) is the systematic study of logic and concerns the *theory* of logic.

It is worth noting that for Peirce there is no *logica utens* "without entertaining a logical theory."

He says:

⁵⁷ My thanks to Alberto Moretti for emphasizing this distinction to me.

⁵⁸ Peirce points to the scholastics as his source (see Pietarinen 2005: 358). He mentions Duns Scotus and Ockham in particular. But the distinction seems to have been known to previous medieval writers, and maybe even before.

The purpose of reasoning is to proceed from the recognition of the truth we already know to the knowledge of novel truth. This we may do by instinct or by a habit of which we are hardly conscious. But the operation is not worthy to be called reasoning unless it be deliberate, critical, self-controlled. In such genuine reasoning we are always conscious of proceeding according to a general rule which we approve. It may not be precisely formulated, but still we do think that all reasoning of that perhaps rather vaguely characterized kind will be safe. This is a doctrine of logic. We never can really reason without entertaining a logical theory. That is called our *logica utens*.” (1903a, CP 4.476)⁵⁹

The requirements that Peirce imposes on the *logica utens* are too strong. He thinks that there is a compulsive instinct or habit that is “logically exactly analogous to inferences,” but acritical and unconscious, and therefore, not genuine reasoning. The possibility of reasoning being deliberate without involving always being “conscious of proceeding according to a general rule which we approve” (however vague the rule might be) is excluded. But if all reasoning is essentially deliberate, fully conscious of an accepted rule, critical, and self-controlled “in the very act of pronouncing the decree,” we risk having to admit that reasoning only happens intermittently, if at all (remember the Tortoise).

We need not follow Peirce here. We can understand the *logica utens* as concerning the act or process of inferring, as the *use* of logical inferences in isolation from the logical principles, rules, or beliefs we may or may not have in mind when performing an inference. The “logic in possession” does not then require explicit acceptance of a rule or principle, but it is often thought to be rule-governed. It will be better for the moment to think of it as proceeding according to rules and

⁵⁹ He also says: “Logic proper is the *critic* of arguments, the pronouncing them to be good or bad. There are, as I am prepared to maintain, operations of the mind which are logically exactly analogous to inferences excepting only that they are unconscious and therefore uncontrollable and therefore not subject to criticism. But that makes all the difference in the world; for *inference* is essentially deliberate, and self-controlled. Any operation which cannot be controlled, any conclusion which is not abandoned, not merely as soon as *criticism* has pronounced against it, but in the very act of pronouncing that decree, is not of the nature of rational inference – is not reasoning. Reasoning as deliberate is essentially critical, and it is idle to criticize as good or bad that which cannot be controlled. Reasoning essentially involves *self-control*; so that the *logica utens* is a particular species of morality.” (1903b, CP 5.108)

remain neutral on whether we follow such rules or we are just conforming to them ('governed' suggests to me the former option, while 'accord' seems more ambiguous on whether the rules guide the practice or we only conform to them).⁶⁰

Logica docens, on the other hand, is concerned with the theory of logic, going from classifications of arguments and characterizations of their correctness, to formal systems and non-classical logics. It is essentially critical and reflective and involves propositional knowledge. It is also, or at least that is the way in which it has been generally understood, essentially normative: it sanctions what inferences may be validly carried out and which can't. We will understand it here in a broad sense, not just as comprising a systematic study of the discipline, the theory of logic (whatever that might be), but also very basic instances of critical reflection about the inferential practice, such as deeming an inference good or bad, appealing to an intuitive notion of validity, and stating basic logical principles.

Now, when it comes to basic logical principles like UI, what Harry lacks is the 'logic in possession,' the *logica utens*. He is unable to perform inferences that, from the *logica docens* point of view, we would classify as instances of the UI pattern. Our attempt to remedy his inferential gap comes, on the other hand, from the *logica docens*. We set out to tell him which logical principle he should accept in the hope that the information given could help him pick up the practice of inferring, the *logica utens*, that he lacks. The adoption problem can be seen as a way of bringing out the issue of the connection between the *logica docens* and the *logica utens*: it challenges the idea that the *logica utens* (at least at a very basic level) requires the (implicit or explicit) acceptance of logical principles or rules (*logica docens*).

⁶⁰ The rules could also be "embodied," that is, hardwired rules that govern basic inferences but are not tacitly known (or "represented") by the subject. (See Devitt 2006: 46ff) This possibility is also left open by the characterization of *logica utens* given above. I leave the discussion of the distinction between following a rule and conforming to it for the next chapter. For 'embodied rules,' see chapter 7, section 7.2.2.

It has been often said that logic as a theoretical discipline could not have arisen until the *logica utens* was at least more or less established. And the point is fairly obvious: the need to, say, classify good and bad arguments could only have arisen once we had something to classify. The same point could be made about other disciplines, like mathematics. A basic mathematical practice had to be in place for mathematical theory to develop. Disciplines like mathematics or logic (both traditionally considered firm candidates for the *a priori*) can, and in fact did, develop to a level of abstraction that goes far past the practice that originated them. But logic seems to have deeper roots in the *logica utens*.

Let's consider for a moment whether it would be possible to adopt a basic mathematical statement. Take a basic mathematical operation like addition and assume we are trying to teach it to someone who has never added in the past. We tell Hanna, Harry's mathematical counterpart, that she should adopt the following axioms: " $(x)(x+0) = x$ and $(x)(y)(x+y) = (x+y)'$." Assuming she is familiar with 0 and the successor operation, she could be able to compute the sum of any two natural numbers by making the appropriate replacements. Now, in principle, there don't seem to be reasons for supposing that this strategy won't work.⁶¹ What we are telling her doesn't require any prior familiarity with addition. In contrast, our attempt to generate a logic in possession by appealing to principles that belong to the *logica docens* dimension has, so far,

⁶¹ The case of counting is more problematic. By counting I have in mind what Benacerraf calls transitive counting, using a progression to count the elements of a set, and not simply intransitive counting, that is, reciting the elements of a progression (see Benacerraf 1965). Psychological studies claim to give evidence that animals and pre-linguistic infants possess numerosity discrimination, which is basically the ability to differentiate groups by size. The relation between this ability and counting a set of objects is not so clear. Learning to apply abstract numerical symbols to numerical quantities does require a lot of training, but even some animals appear to have mastered this. For discussion of these issues see, for example, Dehaene (1997), Gallistel (1990), and Carey (2009). I will set the case of counting aside and won't make any claims as to whether it could be 'adopted' or not.

failed. The AP seems then to put in a different light the importance and preeminence of the practice for the case of logic.⁶²

Our discussion of basic mathematical statements appears to confirm Kripke's view that the impossibility of adopting certain basic logical principles like UI or MPP shows that there is something special about them, something that sets them apart.⁶³ Some other basic statements, also traditional candidates for the *a priori*, such as "red is a color," don't seem to raise the adoption problem either.⁶⁴ But maybe inferring in accordance with UI is more akin to basic

⁶² In the *Meno* Plato introduces his famous 'eristic argument' that in essence holds that what you know you cannot learn since you know it already, while you cannot discover what you do not know because you will not recognize it when you see it. The way out is given by the doctrine of recollection: all learning is the recovery of knowledge from, according to a literal reading of Plato, a previous existence. In the dialogue the doctrine is supposedly illustrated and confirmed by a slave boy who has never learned any geometry, but from whom Socrates elicits some geometrical truths. The experiment is intended to show that the slave must have had the knowledge all along; he is now simply recalling what he already knew. The inference is then drawn that there is never a state of absolute ignorance. The slave knew those truths; he just needed to be reminded of them.

Note that if Socrates's attempt to elicit geometrical truths from the slave had failed, the option of just telling him those truths would have remained open. Socrates could simply tell him the relevant geometrical principles, and if the slave were willing to defer to Socrates's authority, there wouldn't be in principle any reason to doubt that he will be able to apply them in the future. In contrast, in Harry's case, even if he happened to remember the truth of "Universal statements imply each of its instances," he would not be able to use the principle if he didn't already infer in accordance with UI. He wouldn't be better off than when we just tell him the principle. My thanks to Jonathan Adler for suggesting this comparison.

⁶³ Kripke sometimes calls them 'unique,' but let's be a bit more cautious and reserve judgment on the matter of *uniqueness*. We have already raised doubts concerning the case of counting (fn. 61). Moreover, another case likely to be subject to the adoption problem involves the first person pronoun 'I.' In fact, Kripke himself has argued that if someone lacks the concept for 'I,' it won't be possible to teach her its use by telling her an instruction or rule such as Kaplan's semantical rule for 'I': "when a speaker *S* say 'I ...' what he says is true (or true of the possible world he is thinking of at the time, if that is relevant) if and only if *S*..." (quoted by Kripke 2011c: 295, from Kaplan's unpublished manuscript "Meaning as Use") Kripke's point is that the application of the rule presupposes that she already has the concept of herself, that is, the concept expressed by 'I.' In particular, he argues that from the rule alone she wouldn't know when to utter 'I' sentences. It seems to me that there are important aspects in common between this case and basic logical principles (but Kripke doesn't make this connection). It is worth pointing out that Kripke thinks that acquaintance with oneself is necessary for us to possess the I concept.

⁶⁴ What about "bachelors are unmarried men" or "brothers are male siblings"? Assuming familiarity with 'unmarried men' and 'male siblings,' someone lacking the notions of bachelor and brother may be able to adopt them on the basis of the statements above. Their use would depend on the subject's ability to make the appropriate substitutions, which would presuppose familiarity with the law of identity. And this law

skills, like boiling an egg or even riding a bike.⁶⁵ So, what would happen if we were given instructions for very basic skills, like the ones just mentioned? Would we be able to ‘adopt’ the relevant skills? This is an interesting suggestion that will be better approached in connection with the discussion of Ryle’s distinction between knowing that and knowing how. (Chapter 7. In contrast with the examples we have been discussing, this issue doesn’t concern the UI logical principle but the capacity or ability to infer according with it).

Before moving on let me say that what we have been arguing so far should *not* be read as implying that the *logica docens* has no impact on our inferential practice. The claim is not, of course, that basic logical principles can’t ever be taught. But usually when these principles are taught, the goal is to teach an abstract and general principle that the student will recognize because she has already made countless inferences in accordance with it, not to make someone who has never inferred in accordance with them start using them.

In the Princeton seminar Kripke says that such instruction may even benefit someone who only sometimes infers in accordance with the principle we are trying to teach. We can imagine that since the subject has already inferred in accordance with the principle, by leaning on those previous uses and the explicit statement of the UI principle, she *may* be able to start inferring in accordance with UI in a more uniform way. Studies in cognitive psychology seem to support that subjects with formal training are better at performing deductive inferences (see, for example, Stanovich and West 2000) – although one might wonder exactly what is responsible for this improvement. It might well be that training and exposure to examples helped them become less biased to content, for example, but that the explicit formulation of the principles is not directly

seems to me to be another good candidate for the adoption problem. Consider: would telling someone the principle “All A’s are A’s” help him realize that “all lions are lions” is an instance of it? Would someone that lacks the concept of identity be able to identify the latter as having the *same* logical form?

⁶⁵ Michael Devitt made this suggestion to me.

responsible for this improvement: they don't deduce from them their applications. Indeed, UI inferences are so basic that a statement of the principle would appear to be superfluous. In principle, however, an inconsistent UI-inferring subject could, unlike Harry, recognize a particular universal statement as an instance of the UI principle and infer by MPP the instances that follow from it.

CHAPTER 3

Rules, Examples, and the Wittgenstein Connection

3.1 Basic Inferential Rules

It probably already occurred to the reader that what Harry really needs is a rule of inference. After all, deductive inferences are supposed to be governed by rules. If we want Harry to perform an inferential act, to derive a conclusion from a universal statement, a rule of inference would seem to be much better suited to the task than a principle.

The point should have been obvious from our discussion of Carroll's note. One of the main teachings of the note – or at least the teaching most people tended to agree on – was supposed to be that there is a fundamental difference between implications and rules of inference. As we noted, telling Harry the UI principle would not introduce a Tortoise-like situation because we can't even expect him to add the principle as an extra premise (if not done by chance, this would already require inferring in accordance with UI). Yet, merely stating the logical relationship between universal statements and their instances may not be the best idea when our goal is to get him to infer, to detach instances from universal statements. So we may in the end be the ones to blame for Harry's cluelessness: the principle just doesn't tell him what to do with a universal statement. If we truly mean to help Harry out, to get him to produce derivations, we need to provide him with some kind of *instruction* as to *how* to proceed. And a rule of inference appears to be our best choice for carrying out this guiding role.

What would happen if we gave Harry a rule of inference? Would he be able to *adopt* the UI rule? By ‘adopt’ we mean that Harry picks up a way of inferring that accords with the UI-inference rule *on the basis* of his acceptance of the UI rule itself. As before, mere causal connections, where his inferential acts are *not* grounded on his acceptance of the rule, are excluded here. We want him to infer in accordance with UI in virtue of his acceptance of the rule; the reason for his UI-inferring should be given by his acceptance of the rule.

Now, if Harry does indeed accept the UI rule proposed, he should at least show commitment to go by the rule, willingness to take the rule as a guide for his future inferential acts. Our hope is that his acceptance of the rule will be sufficient to prompt him to act *on* the UI rule. In other words, we expect Harry to be able to *follow* the UI rule *because* he has accepted it.

Specifying conditions for rule following is not exactly easy.⁶⁶ To avoid more trouble than we can handle at the moment, we will forget for now one of the problems commonly associated with Wittgenstein’s rule following considerations (and Kripke’s *WRPL*): we will assume that we are able to ‘read off’ which rule a subject is following from her inferential acts (see section 3.3.1 for discussion of this problem). Moreover, rather than thinking of rule following in a general way, we will concentrate on the options that could be relevant to Harry’s peculiar situation.

⁶⁶ It hardly needs to be pointed out that the discussion of this topic originates with Wittgenstein. In the *Remarks on the Foundations of Mathematics*, he says:

[...] “How can one follow a rule?” That is what I should like to ask.

But how does it come about that I want to ask that, when after all I find no kind of difficulty in following a rule?

[...]

How can I follow a rule, when after all whatever I do can be interpreted as following it?

What must I know, in order to be able to obey the order? Is there some *knowledge*, which makes the rule followable only in *this* way? [...] (Section VI, §38)

In Kripke’s famous reading, Wittgenstein’s discussion of rule following is taken to introduce a skeptical problem. *WRPL* does not provide, however, a characterization of rule following. For characterizations of rule following see, for example, Boghossian (2008a) and Pettit (1990a and 1990b).

From what we have already said it is clear that sheer conformity to the UI rule should be distinguished from UI rule following. Harry could just happen to conform to UI by chance, or because of some other reason which is *not* based on our telling him the rule and his acceptance of it (as in the case proposed by Hempel). Moreover, the fact that a rule can be used to accurately describe and predict a subject's behavior doesn't warrant the assumption that such rule constitutes a guide *for him*. He could be behaving *as if* he were following the rule but not actually going by it (even assuming that we are able to 'read off' a single rule from his inferential acts), and we would thus just have conformity. The possibility of mistakes or failures to act under the appropriate circumstances is also always very much in the cards. The subject's actions or responses won't match, in such cases, the constraints imposed by the rule, and hence we won't have conformity. Nevertheless, he could still be following the rule in question, or at least attempting to do so. Though, of course, the more random the actions or replies, the more we will doubt that the subject is following the rule, even when he thinks he is (as Wittgenstein points out, "... And to think one is following a rule is not to follow a rule ..." *PI* §202).

Attributions of rule following as opposed to rule conformity can be difficult, but we won't worry about this. Our priority here is to get Harry to infer. Only if he does, would the question of rule following versus rule conforming arise. At the moment we can just assume that if he has accepted the UI rule and is conforming to it, he is indeed following it.

Now, in the case of deductive inferences, it would be implausible to require an explicit access to the content of the rule. Most thinkers are not be aware that their inferences accord with rules of inference, let alone are able to state them. Even when confronted with explicit statements, it may be hard for them to recognize that some of the inferences they perform accord with such formulations. Indeed, if such requirement were upheld, deductive rule followers would be rare.

If not explicit, the question then turns on how to construe rule following for deductive inferences. Let me briefly outline and reject three main possibilities. One is to make the acceptance of the basic rules of inference implicit or tacit. On this view, UI inferring is grounded on the implicit acceptance of the rule, understood as a mental state encoding the UI rule. Inferential acts are then guided by the rule, though the mental state encoding the rule is not directly accessible to the inferring subject. A second possibility is that deductive inferential rules are ‘hardwired’ in us and govern our inferential acts without being represented.⁶⁷ When we infer, we go through a processing operation that simply pairs an input with an output. The process itself abides by, say, the UI rule, though not because it encodes the content of the rule, but because the UI rule itself is hardwired in it. Even though it doesn’t make much sense to talk of a processing operation as being guided by the UI rule, we can say that we are governed by the UI rule (it seems that neither us nor the processing operation would be ‘following’ the rule given the lack of content). We could, in any case, still say that it is *because* the UI rule is hardwired in us that we infer conforming to it, but ‘because’ would have to be understood as merely causal. Finally, the third possibility is that basic inferential acts are not grounded on rules, whether (explicitly or implicitly) accepted or hardwired. On this view inferential acts are a primitive ability or capacity, which sets in motion regularities or patterns of use that constitute the basis of our *talk* of rule following. Though a direct involvement in shared inferential practices has an impact on the patterns we favor, the inferential acts are not – at least at the most basic level – internally informed by inferential rules. In such cases, logical principles and rules can provide guidance, but

⁶⁷ Devitt defends a version of this view (Devitt 2006, 2011a, and 2011c). He distinguishes between “rules that govern by being represented and applied from ones that are simply embodied without being represented.” (2006: 46) See chapter 7, section 7.2.2 for discussion.

their role is extrinsic to the inferential acts: while the rules of inference are not constitutive of the inferential acts themselves, they can be used to evaluate the inferences resulting from such acts.

It should be clear that none of these proposals would be appropriate for Harry. Harry, who has never inferred in accordance with UI, has presumably not accepted the UI rule implicitly. But whether we could get him to do so or not should not be a reason for concern. The point of construing acceptance as implicit was to give thinkers a break: explicit acceptance was waived because it was deemed too demanding. We are, however, perfectly able to explicitly state the rule for Harry. And if an implicit acceptance of UI were to be sufficient – if it were because a thinker has accepted the rule UI that she is able to perform UI inferences – there is no reason to suppose that in Harry’s case, once the rule has been stated for him, an explicit acceptance would do worse. On the contrary, one might think that Harry should be in a better position to perform UI inferences than someone who has only implicitly accepted the rule.⁶⁸

The second and third proposals are not of much help either. There isn’t much we can do if Harry just happens not to embody basic inferential rules or have a primitive inferential capacity. Were we still willing to help him out, we should be thinking along completely different lines (brain surgery?), as we are certainly not going to be able to *talk* him into inferring. These proposals bypass the need to appeal to an (implicit or explicit) acceptance of the inferential rules, and thus already presuppose that they cannot or don’t have to be adopted. For both accounts, the rule’s propositional content is superfluous to the inferential act. The conceptions of the *logica utens* implied by these views are at odds with Peirce’s characterization (chapter 2, section 2.2.2), who thought that only reasoning that is deliberate, critical, and self-controlled is worthy of being

⁶⁸ Note that if implicit or tacit acceptance of the rule is not understood in terms of a mental state encoding the content of the rule but in terms of inferential regularities, then the first option collapses into the third.

called reasoning. We will deal with them in chapter 7. But for now we must remain (cautiously) optimistic on the matter of adoption.

Let's then stick to an explicit formulation. In the Princeton Seminar, Kripke proposes the following rule:

From every universal statement, deduce each instance.

As Boghossian (2008a: 111) points out, when Kripke talks about rules, he usually has imperatives in mind, and this case is no exception. They seem to be Wittgenstein's favorites as well. There is, however, an immediate problem with this rule: if Harry were to follow it to the letter, we would have him inferring each and every one of the instances of every universal statement he comes across. And one may well think that this would be either worse than not inferring at all (in the cases where the domain is finite), or simply impossible.⁶⁹

Following a suggestion of Boghossian, we can replace it by a norm of permission:

From any universal statement, you may deduce any instance.⁷⁰

⁶⁹ Boghossian (2008a: 111ff) argues that one of the main problems with imperatives is that, while they can easily account for requirements, they have trouble accommodating permissions. Norms, such as the one proposed immediately below, can be easily construed in either terms of permissions or requirements. He identifies, however, a circularity problem when norms are construed as epistemic norms (Boghossian 2008a: 114-115).

⁷⁰ In view of this, Kripke proposed to me the following alternative formulation: 'When confronted with a universal statement and an instance of it, always infer the latter from the former.' Though this formulation appears to avoid the problem of having to deduce each instance of every universal statement, it still requires an inference every time we come across a universal statement and an inference. Moreover, its antecedent also requires identifying a universal statement *and* an instance of it, which would seem to generate an additional problem for someone like Harry.

Note that the rule could also be formulated as an epistemic rule in the manner suggested by Boghossian (2008a: 109). For UI we would have something like this: For all universal statements, if you are rationally permitted to believe it, then you are rationally permitted to believe that a particular instance holds (or, for

Here inferences are permitted but not required. In Boghossian's words, such norms state that, "if a certain condition were satisfied, then performing some particular action would be alright." (p. 112) Though this formulation helps with our previous worries, it is likely to complicate matters even further. Should Harry deduce any instance at all? What circumstances should he take into account for deciding? The imperative's lack of subtlety, on the other hand, may make things easier for him. More importantly, the permission formulation appears to be closer to the principle we tried before, and we all know how that worked out. We'd better let Harry have the imperative and worry later when (and if) he gets stuck with irrelevant inferences.

The rule then demands a certain course of action in connection with a universal statement and Harry is asked to accept it and go by it. Suppose Harry is willing to do this and even reformulates the rule as a personal imperative: "from every universal statement I am always to deduce each instance."⁷¹ We wanted him to tell us whether Alex the lion talks, and we know that he knows from the review that all animals in the *Madagascar* movie talk, and that Alex the lion is an animal.

How would he go about in applying the rule? On the one side we have the inference we want him to make:

All animals in the movie *Madagascar* talk

Alex the lion is an animal in the movie *Madagascar*

Alex the lion talks

all P , if you are rationally permitted to believe $(\lambda)Px$, then you are *prima facie* rationally permitted to believe Pa). These variations as to how to construe the rule won't affect the argument.

⁷¹ Those who remember *WRPL* will find this move familiar.

And on the other hand we have the rule he has accepted:

From every universal statement, I am always to deduce each instance.

Now, to begin with, if the rule is to be of any help to Harry, of any guidance, he needs to recognize “all animals in the movie *Madagascar* talk” as a universal statement. Otherwise, he won’t know that the rule can be applied in *this* case. Recognizing the situations in which a rule applies is obviously crucial if we are going to follow the rule. Unfortunately, in this case the identification of the conditions in question happens to be itself a situation that calls for an application of the UI rule.

Let’s paraphrase the UI rule as follows:

For all x , if x is a universal statement, then, for all y , if y is an instance of x , then y follows from x .

$$[\forall x(\mathbf{U}x \rightarrow \forall y(\mathbf{I}yx \rightarrow \mathbf{F}y))]$$

The rule is then a universal conditional statement with an embedded universal conditional statement in the consequent. As we said, if the rule is to instruct Harry, he first needs to establish that a particular statement satisfies the antecedent of the UI rule, which is itself a universal statement. Of course, to get to the conclusion we want, he will also need to recognize a particular instance of the previous universal statement as satisfying the antecedent of the second conditional

(so even if we told him that this statement in particular is a universal statement, it wouldn't help him). But it is pointless to worry about this last part; Harry will never get there. The reason is fairly obvious, on the intended reading: the only way of realizing that the antecedent of the main conditional obtains, is by performing a universal instantiation. And, as we are assuming, he has no idea how to do that. Since he won't notice that the initial condition of the rule (the antecedent) obtains, the rule itself will be useless from Harry's perspective. So Harry is faced with basically the same problem he encountered before and no progress has been made by switching to rules.⁷²

The difficulty seems to be that the rule requires an inferential connection to its applications. Because of this, for the rule to tell us what to do, it has in effect to function as a sort of extra premise of the particular inferences it is supposed to apply to.⁷³ Contrary to what we may have thought, once the rule is accepted or internalized, it does not just operate on its own, executing inferences while we simply obey it without having much of a say. The rule requires an interpretation: we need to establish that the conditions specified in the antecedent of the rule obtain, only then we can turn to whatever the rule is telling us to do. In the case of a basic case

⁷² Kripke's immediate reaction to the idea that adoption may be possible by switching to rules is that it won't make any difference. He says:

... you may think, 'Ah, the point of this is that we shouldn't adopt a statement ... we really mean to adopt a rule. But what does 'a rule' mean? If it is an imperative that I tell this guy, that will not help him either.

Suppose he agrees to obey. I, of course, will not regard him as having really obeyed if then he doesn't infer universal statements and their instances. But if we are adopting a so-called neutral standpoint from which it is supposed to be up for grabs what follows until we are told what rule to adopt, then you can quite well say, 'All right, I will always infer each instance of a universal statement from the universal statement itself. Now, I don't see that that means, though, that I must infer this particular instance of this universal statement.'

Without begging any questions, I don't see that that follows either. After all, if the inference wasn't obvious to him in standard declarative logic it will not be obvious in imperative logic either. So the adoption of something that he will assent to – of this statement as a rule – will not help him either.

⁷³ I say 'sort of' because in this case the rule is an imperative and thus not truth apt.

like UI, establishing the connection between the rule and its applications becomes very tricky: no correct identification of the initial condition of the rule is possible – and consequently, no UI rule following is possible – without a prior inference in accordance with the UI rule.⁷⁴

Not only the interpretation could not happen at all, it could also go very wrong. To illustrate the point, consider Harry's cousin, Ernie. Like Harry, Ernie has never reasoned in accordance with UI, but has long been using a very different pattern of inference, the pattern of 'Perverse Instantiation' (PInst.).⁷⁵ According to PInst., all universal statements imply the negation of all of their instances. So, from "For all x , $P(x)$ ", "not $P(a)$ " follows.⁷⁶ Ernie does not explicitly state PInst., but we find him always inferring in accordance with it. One day, while watching *Madagascar*, we hear him mutter to himself, "*all* lions are vegetarians, so Alex the lion is *not* a vegetarian." As the story unfolds and it becomes clear that Alex is indeed a carnivore, looking very pleased with himself, Ernie exclaims: "of course! All lions are vegetarians, therefore Alex the

⁷⁴ For MPP the problem seems to be that for a MPP imperative rule to tell us what to do, we will need to detach the consequent of the rule, but that requires inferring in accordance with the MPP rule (since the rule itself will also be a conditional, containing the demand in the consequent).

MPP is needed, however, for any rule because of their conditional form. Wright (2007: 491ff) calls this the "modus ponens model of rule-following." See also Boghossian (2008a) for discussion of the problem this point generates in connection with inferences, and section 3.3.2 below.

⁷⁵ Kripke discusses the case of perverse instantiation and in the Princeton Seminar he attributes the name to Kit Fine. He uses it to illustrate the problem of determining what meta-logic would be appropriate for comparing alternative logical principles. In Ernie's case, because it will not be possible to agree with him on what follows from a universal statement, the idea is that it is difficult to find a neutral ground for evaluating the two principles, UI and PInst. I am using it here in another way.

⁷⁶ It might be argued that, if the principle of charity were to be believed, we wouldn't be able to identify Ernie as a perverse instantiator. Quine and Davidson have argued that basic logical truths are imposed on our interpretation of other speakers. In Quine's view, "illogical cultures are indistinguishable from ill-translated ones" (1954: 105. See also fn. 111 and 113 below and Quine (1960: 57 and 68) for remarks on the myth of a "prelogical mentality"). In these lectures Kripke raises some objections to the principle of charity (though not in connection with this), where he basically claims that the circumstances under which an exchange takes place need to be taken into account, and argues that a model as rigid as the one offered by Quine would clearly deliver erroneous results (he discusses the example of a very primitive group of people translated as believing that the Earth goes around the Sun). Setting this aside, and admitting that it is likely that we would find Ernie's performances somewhat puzzling at the beginning, I can't see any good reason for claiming that we won't be able to figure out that he is using P.Inst.

lion is not a vegetarian!” Suppose that at this point and against our better judgment we decide to intervene and tell him that the pattern of inference he has been following is bad one. We say to him, “look Ernie, all principles of the form ‘all universal statements imply the negation of their instances’ are fallacious. Please stop using them. Remember: if any inference has the form ‘all universal statements imply the negation of their instances,’ conclude that it is fallacious!” Let’s assume he believes us and even explicitly acknowledges that he accepts the principle that all inferences of the form ‘all universal statements imply the negation of their instances’ are fallacious. We then tell him that ‘all lions are vegetarians, therefore Alex the lion is not a vegetarian’ is fallacious. But instead of assenting, he says: “No, therefore, it is not fallacious!”

Neither Harry nor Ernie have benefited from our interventions. Harry could not identify the initial conditions of the rule, and Ernie interpreted it in his usual way, according to PInst. So once again we have failed. Basic rules of inference cannot be *adopted*, and as Ernie’s example suggests, dropping them can be difficult as well.

Before moving on, let me anticipate a very natural complaint that has already been raised by Black (1970) in connection with Carroll’s note: neither Harry nor Ernie have either accepted or understood the rules we gave them. I think the objection, though correct, misses the point. In Harry’s case, the relevant point is that he *could not* have followed the rule. From Harry’s point of view, if rule following is to be based on acceptance, UI rule following is impossible.

Putting the matter in its most perplexing form we can say this: Harry will not be able to follow the UI rule without already being able to perform UI inferences; to begin following the UI rule he has to conform the UI rule already.

3.2 Examples: The Last Attempt

Another rejoinder to what we have been saying may be the following: who could possibly think that giving Harry a rule of inference, all by itself, would be sufficient? As anyone with even minimal pedagogical sense would know, some examples illustrating the application of the rule are needed as well!

The objection seems perfectly reasonable. Perhaps a set of examples will give Harry a grip on how the rule is *meant* to be applied. But how exactly is a set of examples supposed to achieve this? It should be clear that the examples are *not* intended to bypass the rule, but to provide an illustration of its applications.

Of course, it isn't hard to imagine that given the right set up and a continuous exposure to particular UI inferences, Harry could be lead to grasp the UI rule by extrapolation (see Kusch 2006: 25). For all we know such exposure could cause him to develop an inferential pattern that accords with the UI rule. Naturally, we don't have a guarantee that this will be the case. As Wittgenstein emphasized,⁷⁷ given that the examples will necessarily be finite and will require an interpretation, he could also end up picking up or grasping some other rule.

These are not, however, the kind of cases we are interested in at the moment. To repeat: we want to see whether Harry could 'adopt' the rule, and acceptance of the rule is essential to our

⁷⁷ About this, Wittgenstein says:

210. "But do you really explain to the other person what you yourself understand? Don't you leave it to him to *guess* the essential thing? You give him examples – but he has to guess their drift, to guess your intention." – Every explanation which I can give myself I give to him too. – "He guesses what I mean" would amount to: "various interpretations of my explanation come to his mind, and he picks one of them". So in this case he could ask; and I could and would answer him.
(PI)

understanding of adoption. Borrowing a distinction from cognitive psychology,⁷⁸ we can say that Harry's adoption of the UI rule is not meant to be a case of implicit learning, that is, a bottom-up process where the rule could be picked up by observation and practice alone. Instead, it is designed to be a case of explicit learning, a top down process that originates from propositional knowledge⁷⁹ and results in picking up a way of inferring according to UI. For this reason, because we want to see if there is any way to get Harry to 'adopt' the rule, the examples have only a heuristic role, and the rule remains the central part of what we tell him. [Recall that we are trying to challenge accounts of inferential acts based on acceptance. See chapter 2, section 2.2.2]

To avoid confusion, let's assume that Harry has been exposed to examples of the UI rule and that such exposure has not resulted in implicit knowledge. We then propose the rule to Harry and follow it with a number of examples. Would this help? Would Harry realize that what we are telling him is not simply a succession of different things with no connection with each other?

It seems clear that for that to happen he will need to see the examples *as examples of* the UI rule. We find, once again, that there is no immediate way of connecting the rule with its applications: for the examples to be interpreted as instances of the rule in an appropriate way, Harry will have to infer in accordance with the UI rule.

⁷⁸ Devitt expresses the distinction between explicit and implicit learning as follows:

Explicit learning is a "top-down" process that starts from declarative knowledge. Consider, for example, learning to change gears in a stick-shift car by starting with instructions like: "First, take your foot off the accelerator, then disengage the clutch." In contrast, implicit learning is a "bottom-up" process: we observe, practice, and "just pick the skill up." (Devitt 2011a: 209)

For more on this distinction and the notion of skill, see chapter 7, section 7.2.2.

⁷⁹ Cognitive psychologists and neuroscientists would replace 'propositional knowledge' by 'declarative knowledge,' though they are commonly treated as equivalent. This identification has recently been challenged by Stanley and Krakauer (2013).

What about MPP? Would the examples help with this case? Suppose we tell him that:

(1) For any A and B, “if A then B, and A, B follows,”

and then we give him some examples like,

(2) “If Alex is a lion then Alex eats meat, Alex is a lion, therefore Alex eats meat.”

Assuming he has no problem with UI, the inference could be as follows: if (2) has the same logical form as (1), then (2) is an example of (1); (2) has the same logical form than (1), therefore ... (we could also infer in an analogous way for each of the premises and the conclusion).

The complaint that we hadn’t given Harry any example seemed reasonable at first, but maybe because we were mostly thinking by analogy with our own case. Basic rules tend to generate immediate recognition among those that infer in accordance with them. Examples are usually given together with the rule for extra help, as a way of illustrating with concrete cases something we have done countless times in the past. Because both the rule and the examples tend to generate recognition, the relation between the two is likely to appear as immediate. Yet the situation would obviously be very different if we had never made those inferences before.⁸⁰

⁸⁰ Notice that in this case, even if it were possible to provide Harry with a list of all the cases of UI for him to memorize, he wouldn’t see them as instances of the UI rule. For that, an inferential relation between the cases and the rule is needed.

3.3 The Wittgenstein Connection

We have already alluded to points of connection between the adoption problem, Wittgenstein's rule following considerations, and Kripke's own *WRPL*. In fact, at the beginning of the Princeton Seminar, Kripke explicitly highlights such connection. He says:

In this way it *is* clear to me that the central point that I might want to talk about here is really intimately related, at any rate, to the central point of Wittgenstein's later philosophy, at least the central point as I see it: the whole ensemble of issues, the argument against private language, properly understood...

And later on:

... one couldn't think of the laws of logic as based on a set of assumptions, whatever else one thought of them as being. [...] I don't know how far this kind of consideration goes in the history of philosophy [...] but it is in the article by Lewis Carroll, who was taken over in one article by Quine but seems to be ignored by him in others. [...] A very similar consideration is basic to Wittgenstein in his analysis of following a rule and mathematical proof, and the impossibility of a private sensation language. All of these are based on a very similar consideration to what I want to talk about. But I don't know if this kind of consideration goes back earlier, and it's hard to say in fact, in the Lewis Carroll case, what kind of moral he thought he was drawing because he just sets up a puzzle.

The announced topic of the seminar was Wittgenstein's *Philosophical Investigations* (the material that later became *WRPL*), but Kripke decided to begin with some considerations on the nature of logic with the intention of getting to Wittgenstein later on. He never did, and the seminar was completely devoted to issues on the epistemology of logic. So he never said what the connection was supposed to be! It is clear, however, that at that point he thought that these two themes,

what we have called the adoption problem and Wittgenstein's rule following considerations, were somehow related. It will be up to us now to try to figure out this relation. Let's begin with a summary of the main argument of *WRPL*.

3.3.1 *WRPL* and the Adoption Problem

Kripke's story begins with someone, let's call him Smith, who has been computing sums of numbers less than 57 satisfactorily, but has never before dealt with numbers equal to or greater than 57. Nevertheless, when confronted with the computation of $68 + 57$ he is certain that it ought to yield the result 125. At this point, a skeptic comes along and suggests that Smith should respond with 5: he insists that by 'plus' Smith has always meant '*quus*,' an operation compatible with addition for numbers below 57 and resulting in 5 for numbers above that threshold, the '*quum*' of the numbers ('*quus*' is symbolized by ' \oplus ': $x \oplus y = x + y$ if $x, y < 57$ & $= 5$ otherwise).

Smith, like any good mortal, has of course only added finitely often in the past, and the difficulty is that it is not possible to uniquely determine from some finite number of uses of an expression the rule applied. Smith's uses of '+' in the past do not *determine* that he should obtain '125' from ' $68 + 57$ ' because those uses are also compatible with his meaning '*quus*' rather than 'plus.' There seems, then, to be no way of determining which rule he is following on the basis of his past uses, and thus there is no justification for his present application of the rule and his claim that by '+' he means 'plus.'

The search for a fact in Smith's own mind which would constitute his meaning 'plus' rather than '*quus*' also turns out to be unsuccessful since neither intentions nor dispositions seem to be of

any help⁸¹ (even though Smith is allowed to have perfect recall and access to all aspects of his former behavior and mental life). This leads the skeptic to conclude that there is no fact of the matter about Smith's meaning one thing rather than another.

It is important to stress that, though introduced in an epistemological guise, the problem raised is metaphysical. Kripke says:

It is clear that the skeptical challenge is not really an epistemological one. It purports to show that nothing in my mental history of past behavior – not even what an omniscient God would know – could establish whether I meant plus or *quus*. But then it appears to follow that there was no *fact* about me that constituted my having meant plus rather than *quus*. (*WRPL*: 21)

Soames outlines the main points of the argument as follows:

P₁ If in the past there was a fact about what I meant by '+,' in particular, if there was a fact that I meant addition by '+,' then, either:

- (i) this fact was determined by nonintentional facts of such and such kinds – facts about my past calculations using '+,' the rules or algorithms I followed in doing calculations involving '+,' my past dispositions to respond to questions 'What is n+m?,' the totality of my past dispositions to verbal behaviour involving '+,' etc.
- (ii) the fact that I meant addition by '+' was a primitive fact, not determined by non-intentional facts.

P₂ Nonintentional facts of type (i) did not determine that I meant addition (or anything else) by '+.'

P₃ What I meant by '+' was not a primitive fact. [Though recall that Kripke says that "such a move may in a sense be irrefutable, and if it is taken in an appropriate way Wittgenstein may even accept it." (*WRPL*: 51.)]

⁸¹ Many, of course, would disagree with this point. See, for example, Soames (1998a, 1998b) and Horwich (1995, 1998) for a defense of the dispositionalist view, and Wright (1984) and Pettit (1990a, 1990b) for a defense of the intentional view.

C₁ Thus, in the past there was no fact that I meant addition (or anything else) by ‘+.’

C₂ By parity of reasoning, there never was a fact about what I, or anyone else, meant by any word; *ditto* for the present. (Soames 1998: 232)

Now, an obvious point of connection between *WRPL* and the adoption problem is that both raise objections to the possibility of rule following. But is that all? Let’s start with some differences.

In Smith’s case, the rule following problem – that is, (at least initially) determining which rule is being followed – is raised in parallel to the question of what determines the meaning of addition. Harry’s case is not, however, explicitly raised in connection with meanings or concepts. We are told that he doesn’t infer in accordance with UI, not that he doesn’t understand the concept ‘all’ or know its meaning. We have seen that it is hard not to think of his failure to infer in accordance with UI as due to his failure to understand ‘all,’ and hence, that whatever he might himself claim, he has not really accepted the UI rule or principle. But this reaction establishes a very tight connection among the meanings or concepts of logical terms, our understanding of them, and their inferential uses. Yet, such connection needs to be argued for and clarified in terms of priorities (see chapter 6).

Kripke has said in conversation that he doesn’t think of Carroll-like cases and the adoption problem as being about the meanings of the logical words. Of course, that doesn’t entail that it could not be thought in that way. But it does reflect the difference between the two cases, since Smith’s problem is explicitly set up in connection with meanings – leading Boghossian to say that, “‘the rule-following considerations’ is, strictly speaking, a misnomer for the discussion on offer.” (1989: 18) The main point is that while for Harry the content of the rule is given and

allowed to remain constant, in Smith's case the meaning of addition and the rule being followed are what has to be determined.

Another difference is that Smith's troubles do not begin with his being a first time adder, but with his having added only a finite number of times. Harry, on the other hand, is someone who has never performed a universal instantiation. In Smith's case the impossibility of determining the rule he is following leaves him without any justification for his present application of the rule, his reply is not justified (or, as Boghossian (2008a: 118) prefers to say, "rationalized," since he may be following bad rules) by a rule. His present course of action is, Kripke suggests, just a "leap in the dark." For Harry, there is no present course of action: the problem is not how to account for his use of the rule or principle, his problem is rather how to use it *at all*. So, we could say that Harry's problem precedes Smith's problem. One could easily imagine that if Harry ever gets to universally instantiate, he could then have a Smith-like problem and a skeptic could wonder whether he is following UI or some other rule UI*.⁸² But he is not there yet.

The point is fairly obvious once inferential acts are understood in terms of rule following: clearly, if there is a paradox arising from the determinacy of the rule being followed, Harry is going to get into trouble again if he ever begins to follow the UI rule. In a way, the situations are reversed: Smith claims to have produced an application of a rule he is following and needs to determine the rule itself to account for his use, while Harry has been given the rule and is asked to produce an application.

A point we already made in our comparison between Harry and the Tortoise can also be appreciated here. The adoption problem is not about how our beliefs in basic logical principles

⁸² I am stressing here the epistemological side of the paradox, which is the aspect Kripke explores at the beginning, when he introduces Smith's case. But, as we saw before, the conclusion he ultimately reaches is metaphysical: there is no fact of the matter regarding what a speaker means in uttering a given word or sentence.

or rules are to be justified, or how we can convince a skeptic of their good standing. And it is not about the determinacy of the rule either. I see the adoption problem – that is, the impossibility of picking up a basic inferential practice by means of the acceptance of the corresponding basic logical principle or rule of inference – as a way of bringing up the issue of the role that these principles or rules have in the inferential acts that accord with them. As we shall see in the following chapters, it will have consequences for a range of different views on the justification of logical principles. But it is conceptually separable, and comes prior to the justification problem and the determinacy problem.

From all this, it would seem that is Harry the one dealing with the worst share of the problem: for him, the UI rule is unfollowable if not observed already; it fails to provide any guidance for those who don't infer in accordance with it already.⁸³ Three potential problems seem to lay in wait for him:

1. How to start performing UI inferences on the basis of the rule he was given.
2. If he ever gets to perform UI inferences, how to justify or rationalize his application of the rule by reference to UI and not some deviant UI* rule.
3. If he manages to determine the rule he is following, he would then face the problem of how to justify the rule itself, that is, the classic problem of giving an epistemic justification of the rule.

The future does not look very bright for those that, like Harry, are trying to base inferential acts on rule following (at least if the latter is construed in terms of acceptance). Indeed, it is hard

⁸³ Notice that the problem also arises for Ernie, our perverse instantiator. It seems to affect basic acts of inferring disregarding questions of validity.

not to wonder what the role of a rule that cannot be followed if not followed already could possibly be. How can they be of help?

3.3.2 Rethinking the Connection

Given these differences, we might think that the adoption problem and the skeptical paradox Kripke introduces in *WRPL* don't have that much in common. Other than being concerned with rule following, they would appear to be addressing different aspects of it. Why, though, does Kripke say that they are "intimately related" and "based on very similar considerations"? At one point he even says this:

... if we couldn't deduce from "All things are black" "The Eiffel Tower is black" without recognizing the truth of the law of universal instantiation, then one won't be able to deduce it even if one does recognize the truth of the law of universal instantiation. Recognizing this truth will not help one unless one was already someone who made inferences of this kind.

OK, I hope this point is clear. This is the basic point. This is the basic point in Wittgenstein's later philosophy. He wasn't talking about logic. This is the basic kind of point. I mean, eventually one can't just give oneself a belief or a premise or something; one just *does* draw certain conclusions. But as it arises in Wittgenstein – that I will talk about later. Here, how can we talk about it here? [Princeton Seminar]

And he goes back to the previous issue just when we thought he was going to say what the connection is supposed to be! But the important point here is that he claims that this is also "the basic point in Wittgenstein's later philosophy," suggesting that what we said before cannot be the whole of the matter.

Now, Boghossian and Wright have argued⁸⁴ that Kripke's interpretation overlooks a point Wittgenstein makes regarding rule following: that rule following – at least when it is understood as an intentional state – is mediated by inference. In their view, there are two aspects to Wittgenstein's treatment of rule following and his regress of interpretations. One of them, the one emphasized in *WRPL*, concerns the determinacy of the content of the rule. The other aspect stresses that following the demands of a rule requires a MPP inference.⁸⁵ Both Wright and Boghossian think that the two aspects are critically different, and that the correct reading of Wittgenstein's regress of interpretations argument is given by the second aspect.

What's exactly the problem this second aspect generates? Boghossian has argued that if we assume an intentional view of rule following, where rule following is explained and 'rationalized' by means of an intentional state that encodes the rule, a vicious regress would seem to be unavoidable. Consider the following example he proposes:

(Email Rule) Answer any email that calls for an answer immediately upon receipt!

Now, what would it take to follow such a rule? Boghossian says:

Well, I have grasped the rule, and so am aware of its requirements. It calls on me

⁸⁴ See Boghossian (2008a: 126ff, 2014: 12ff) and Wright (2001b: 208ff; 2007: 490ff; 2012: 380ff).

⁸⁵ Wright (2012: 380) distinguishes the two aspects by means of the following questions: "How can I tell which rule I used to follow? (Or: how is the rule I used to be identified?)," and "How can I tell what the rule I grasp requires of me here?" It is the first question the one that he takes to be prominent in *WRPL* (note how Wright is emphasizing here the epistemological side of the paradox). The second question introduces what Wright calls the "Modus Ponens model of rule following."

It should be stressed that though Wright characterizes a way of conceiving rule following as the Modus Ponens model of rule following in his earlier papers (2001b and 2007), it is Boghossian (2008a) who makes the point that, if inferences themselves are understood in terms of rule following, rule following cannot be thought as involving inferences without setting off a vicious regress.

to answer any email that I receive immediately. I am aware of having received an email and so recognize that the antecedent of the rule has been satisfied. I know that the rule requires me to answer any email immediately and so conclude that I shall answer this one immediately.

On this Intentional construal of rule-following, then, my actively applying a rule can only be understood as a matter of my grasping what the rule requires, forming a view to the effect that its trigger conditions are satisfied, and drawing the conclusion that I must now perform the act required by its consequent. In other words, on the Intentional view of rule-following, rule-following requires inference. (2014: 13)

The problem applies, Boghossian notes, to all rules, “good or bad, justified or unjustified”:

Since any rule has general content, if our acceptance of a rule is pictured as involving its representation by a mental state of ours, an inference will always be required to determine what action the rule calls for in any particular circumstance. (2008a: 127)

And of course, because inferences are supposed to be “a form of rule following par excellence,” we have gotten ourselves in a big mess:

... applying the Email Rule requires, as we have seen, having an intention with the rule as its content and inferring from it a certain course of action. However, inference, we have said involves following a rule, in this case, MP*. Now, if the Intention View is correct, then following the rule MP* itself requires having an intention with MP* as its content and inferring from it a certain course of action. And now we would be off on a vicious regress: inference rules whose operation cannot be captured by the intention-based model are presupposed by that model itself. (Boghossian 2008a: 128)

We said that Boghossian claims that it is this problem and not the one emphasized by Kripke in *WRPL* concerning the determinacy of the rules, the one that correctly captures Wittgenstein’s regress of interpretations argument. In his view, Kripke’s interpretation is unable to explain passages such as:

219. “All the steps are really already taken” means: I no longer have any choice. The rule, once stamped with a particular meaning, traces the lines along which it is to be followed through the whole of space. – But if something of this sort really were the case, how would it help me? [...] (*PI*)

According to Boghossian, on Kripke’s reading there is no explanation of why the rule could have a determinate meaning and still not help.⁸⁶

As we know, Kripke’s seminars and lectures on the nature of logic – which, most likely, neither Wright nor Boghossian had an opportunity to read – are evidence that he was well aware of a related problem, that is, the adoption problem. The two problems are not the same – among other things because adoption does not *just* involve MPP – but they both make the point that applying a rule requires inference (at least when the rule in question is a rule of inference). Yet, contrary to Boghossian and Wright’s opinion, Kripke thinks that adoption is intimately connected with *WRPL*’s problem. Indeed, already there, we find him saying that:

⁸⁶ Kripke *does* have a reply to Boghossian’s question. In a footnote, he says:

... although it is useful, following Wittgenstein himself, to *begin* the presentation of the puzzle with the observation that I have thought of only finitely many cases, it appears that in principle this particular ladder can be kicked away. Suppose that I had explicitly thought of *all* cases of the addition table. How can this help me answer the question ‘68+57’? Well, looking back over my own mental records, I find that I gave myself explicit directions. “If you are ever asked about ‘68+57’, reply ‘125!’” Can’t the skeptic say that these directions, too, are to be interpreted in a non-standard way? (See *Remarks on the Foundations of Mathematics*, I, §3: “If I know it *in advance*, what use is this knowledge to me later on? I mean: how do I know what to do with this earlier knowledge when the step is actually taken?”) It would appear that, if finiteness is relevant, it comes more crucially in the fact that “justifications must come to an end somewhere” than in the fact that I think of only finitely many cases of the addition table, even though Wittgenstein stresses both facts. Either fact can be used to develop the skeptical paradox; both are important. (*WRPL*: 34, fn. 52)

The problem is that even if I had thought of all the cases of addition, I would still have to give myself instructions concerning the application to particular cases. Yet, such instructions would also take the form of a rule: “If you are ever asked about ‘68+57’, reply ‘125!’” But then nothing has been gained, because this rule could also be reinterpreted ... and so on.

Wittgenstein regards the fundamental problems of the philosophy of mathematics and of the ‘private language argument’ – the problem of sensation language – as at root identical, stemming from his paradox. The whole of §3 is a succinct and beautiful statement of the Wittgensteinian paradox; indeed the whole initial section of part I of *Remarks on the Foundations of Mathematics* is a development of the problem with special reference to mathematics and logical inference. (*WRPL*: 20)⁸⁷

And if we take a look at §3 in the *Remarks* we find this:

3. *How do I know* that in working out the series + 2 I must write “20004, 20006” and not “20004, 20008”?
 – (The question: “How do I know that this colour is ‘red’?” is similar.)
 “But you surely know for example that you must always write the *same* sequence of numbers in the units: 2, 4, 6, 8, 0, 2, 4, etc.” – Quite true: the problem must already appear in this sequence, and even in *this* one: 2, 2, 2, 2, etc. – For how do I know that I am to write “2” after the five hundredth “2”? i.e. that ‘the same figure’ in that place is “2”? And if I know it *in advance*, what use is this knowledge to me later on? I mean: how do I know what to do with this earlier knowledge when the step actually has to be taken?

The last questions raise the issue that according to Boghossian would be “mystifying” from a Kripkean point of view (2008a: 126). Yet Kripke appears to think that the “two aspects” of the rule following considerations constitute a statement of the paradox he finds in *PI*. And he doesn’t seem to be alone in thinking this, since Wittgenstein himself tends to present them together as

⁸⁷ See fn. 87. In unpublished work Kripke also singles out the following passage as one of his favorite passages:

How can the rule determine what I have to do?
 Following a rule presupposes agreement.
 It is essential to the phenomenon of language that we do not dispute about certain things. (MS 165: 30ff)

well.⁸⁸ So, are these two aspects indeed “critically” different as Boghossian and Wright seem to think?

It seems to me that despite the differences noted in the previous section, Kripke (and perhaps Wittgenstein as well, but who knows) saw the two problems, the one concerning the determinacy of the rule and the adoption problem, as two different ways of arriving at the same conclusion: if rule following is conceived in terms of an instruction that each of us accesses in the privacy of our own minds, then rule following is not possible. If we ‘looked into our minds,’ we would simply find that there is no rule there to be found. And the reason given is that if we had it, it would be of no help: there is no unmediated access to the rule. The rule is always subject to interpretation, either of its content – as emphasized in *WRPL* – or of its demands by means of inference – as stressed by the adoption problem. The rule itself is then an “unnecessary shuffle” (see next section).

Of course Kripke does not quite put things this way, since he is raising the issue mainly in connection with the meaning of addition. But the two issues, determining which rule is being followed and determining what is meant, whether *quus* or plus, are raised in parallel – as passages such as, “when people speak of themselves and others as meaning something by their words, as following rules, they do so with perfect right ...” (*WRPL*: 69) indicate. If a fact constituting one of them were to be found, it would also constitute a solution to the remaining issue. Moreover, we should not forget a point made in the previous section and sometimes not sufficiently highlighted. Though in *WRPL* the problem is first introduced as an epistemological problem regarding

⁸⁸ This point is noted by Wright:

The distinctness of this issue is somewhat obscured in Wittgenstein’s own exposition in those particular paragraphs, where the mediating activity is represented as one of *interpretation*, so that it is the question of the determinacy of the content of the rule that still seems to be in play. Nevertheless the second aspect is different; and crucially so. (Wright 2012: 381)

Smith's justification for an alleged application of the rule, in the end the problem is metaphysical: there is no rule (or not fact about the rule) supporting his applications of the function, or at least none that he has accepted and guides his answers.

Borges once wrote that "there is a concept which corrupts and upsets all others. I refer not to Evil, whose limited realm is that of ethics, I refer to the infinite." ("Avatars of the Tortoise": 202). And though Kripke claims that the fact that we are only able to compute finitely many cases is in the end irrelevant (see fn. 87), it seems to me that Borges is right, and that in both Harry's and Smith's cases the source of the trouble lies in the infinite-finite relation. Yet the problem is not the infinite *per se* but rather how finite minds can relate to it, how can we access the content of a rule, general by nature, and, therefore, potentially infinite in its applications. It would seem that no matter how immediate, almost given, such access might feel, we always have to interpose some kind of interpretation when the time to apply it comes. The problem worsens the more basic the content in question becomes. And if the content in question is as basic as the UI rule, there appears to be no way of interpreting the dictates of the rule without a transition that accords with the UI rule itself.

To sum up, it seems to me that, though approached in a different way, the main question *WRPL* and the adoption problem raise is ultimately the same: how is rule following possible? And the answer is the same as well: it isn't. At least if rule following is construed as mental states encoding the rule that each of us accesses in the privacy of our own minds. And though approached in a different way, the reasons given for this answer also appear to converge: there is no unmediated access to a rule. The relation between a rule and its applications is always mediated by interpretation, either by inference or by the interpretation of the concepts or meanings involved – or, more likely, by both.

A final point. Boghossian interprets Wittgenstein's remark on blindness in the second part of §219⁸⁹ as claiming that rule following cannot be understood as being "always fully sighted, always fully informed by some recognition of the requirements of the rule being followed" (2008a: 130). Here is the rest of §219:

[...] No; my description made sense only if it was to be understood symbolically. I should say: *This is how it strikes me*.
When I follow the rule, I do not choose.
I follow the rule *blindly*.

So in Boghossian's view, Wittgenstein's metaphor of blindness should be understood as holding that at least some rule following must be such that accepting the rule does not consist in a mental state encoding the requirements of the rule.

Now, if what we said above is correct, the passage may afford a more radical reading. Perhaps we do not rely on rules. At least in basic inferential cases they may be superfluous, "an unnecessary shuffle." We *say* that we follow rules, but that doesn't mean that when it comes down to the inferential act itself, we are *acting on* a rule. The next section will hopefully provide some textual support for this interpretation.

3.3.3 Kripke and Wittgenstein on Rules and Inferences

Kripke's dealings with his alter-ego, known to all as Kripkenstein, are complicated. On the one hand, he finds himself in agreement with much of what Kripkenstein says in setting up the skeptical argument, especially in connection to dispositions. On the other hand, he claims not to

⁸⁹ We will see in chapter 6 that this passage and the blindness metaphor play an important role in Boghossian's own proposal for the justification of logical principles.

“believe the argument” – which I take to mean that he does not believe Kripkenstein’s skeptical conclusion – and then proceeds to point to passages in *WRPL* where his own self is supposed to clearly come through (as *WRPL*: 94, fn. 76).

The tension between the two is not unique to *WRPL*; we will have occasion to appreciate that it also emerges in connection with the adoption problem and what is ultimately to be made out of it (chapter 5 and 7).⁹⁰ At the moment, I just want to point to some passages that show considerable agreement with Wittgenstein himself on inferring and the role of rules. What makes the lectures and seminars on the nature of logic particularly interesting is that in this case Kripke is speaking in his own name. [Though perhaps he would now like to claim that certain passages are truly due to his personal version of Mr. Hyde – Kripkenstein himself. See chapter 7, section 7.2.3.]

Now, Wittgenstein’s later work is commonly thought to be conceived in opposition to the views he held in the *Tractatus*. Regarding the role of rules of inference, however, already in the *Tractatus* we find him saying:

5.132 If p follows from q , I can conclude from q to p ; infer p from q . The method of inference is to be understood from the two propositions alone. Only they themselves can justify the inference. Laws of inference, which – as in Frege and Russell – are to justify the conclusions, are senseless and would be *superfluous*. [My emphasis.]

Years later, upon his return to Cambridge in the early 30s, he still held a similar view:

Inference is the transition from one proposition to another, a transition which we

⁹⁰ The main source of such tension is, in my view, due to Kripke’s resistance to Kripkenstein’s communitarian tendencies. There is a secret well kept from the reader in *WRPL*: Kripke’s heart belongs to what in *WRPL* he calls “platonism.”

justify by saying that e.g. q follows from p . This relation is entirely determined when the two propositions are given. It is entirely different from other relations, in which the opposite case is always thinkable. The relation of following and similar relations are internal relations which hold when (roughly) it is unthinkable that they should not hold. Whether a proposition is true or false can only be decided by comparison with reality. So that $p \vee q$ follows from p, q is not a proposition: it has no use. What justifies the inference is seeing the internal relation. No rule of inference is needed to justify the inference, since if it were I would need another rule to justify the rule and that would lead to an infinite regress. We must see the internal relation. (Wittgenstein 1982: 56)⁹¹

This passage seems to indicate that Wittgenstein was well aware of Carroll's note. At least in the 70s, Kripke seemed to have a rather similar view regarding the role of the rules of inference:

I mean look, what would my own standpoint be? What's the naive standpoint about all of this? It's that, well, from p and q , the conjunction follows, independently of what logic you accept. Accepting logic has nothing to do with it. If someone doesn't accept my logic, that doesn't matter, ' p and q ' will follow. He says p and he says q . If he doesn't infer it, that's so much the worse off for him, he hasn't inferred something which really follows. But if you don't like that standpoint, I mean, that would be the old sort of reactionary standpoint. If you don't like that standpoint, it's no good to say, well, from p and q , ' p and q ' provided you accept a standard logic, or if you accept this rule of inference. There is nothing else that you can accept which is going to get you to this conclusion if not already. (Princeton Seminar)

And later on, while complaining about 'the Quinean picture' (see next chapter), he says:

Obviously we see that certain things follow from certain others. Rather than try to think this as mysterious – of course nothing could be less mysterious than that if all things are black, then the Eiffel Tower is bound to be black – recognizing a

⁹¹ Note also the following passage from §3 of *RFM* (the paragraph highlighted by Kripke in *WRPL* and partially quoted before): "[...]“But do you mean to say that the expression ‘+ 2’ leaves you in doubt what you are to do e.g. after 2004?” – No; I answer “2006” without hesitation. But just for that reason it is *superfluous* to suppose that this was determined earlier on. My having no doubt in face of the question does *not* mean that it has been answered in advance. [...]” (*RFM*, Part I. My emphasis.) Paragraphs §185-192 of *PI* also deal with this issue of how one can determine in advance what is *meant* by a given rule.

principle of logic here seems to me to be rather ... you know, an unnecessary shuffle to use Wittgenstein's phrase, a piece of extra baggage. But if one were in a position where one wants to leave that in suspense, then you can leave anything you please in suspense. You can choose to adopt both hypotheses – that all universal statements imply all of their conclusions, and the imperative of telling oneself “From each universal statement I will draw each conclusion” – and still refuse to deduce this conclusion from this premise. Why not? If you were free without the principle of logic, why aren't you free *with* the principle of logic? (Princeton Seminar)

Both Kripke and Wittgenstein emphasize the need to ‘see’ that something follows (they may, however, have different views on what this ‘seeing’ is supposed to be. See chapter 5). They also seem to have in mind the same culprit: the formalization of logic has given us the impression that rules of inference are all that is needed for inferring. Kripke says:

Wittgenstein, in some passages in the *Remarks on the Foundations of Mathematics*, actually says that formal logic has completely perverted the thinking of philosophers, or something like that. I don't really quite hold to such a statement. But I think it does apply, so to speak, here. What has given people the very misleading impression that there can be alternative logics in the same way there are alternative geometries is the existence of alternative formal systems. You see, one might think, “Look, logic's got to be different from geometry, because we need to use logic to draw conclusions from any hypothesis whatsoever.” Well, now, this is supposed to have been answered with the notion of a formal system. Formal systems can be followed blindly; a machine can check the proof; you don't have to do anything at all; etcetera, etcetera. Nothing – and this will be a theme when I discuss Putnam in particular – nothing is more erroneous. Of course, a machine can be programmed to check the proof. But we are not machines. [...] But anyway, we are not machines in the sense of not using reasoning to understand what we are told, what formal system this is. We are given a set of directions, that is, any statement of the following form is an axiom: if these premises are accepted any conclusion of this form must be accepted. These directions themselves use logical particles, as Quine has rightly pointed out. Because of this use of logical particles, understanding what follows in a certain formal system itself presupposes a certain understanding of logic in advance, and cannot be done blindly.⁹² (Princeton Seminar)

⁹² Kripke is reversing the blindness metaphor here: we can't follow rules without an interpretation, without reasoning about their instructions. We cannot simply obey them blindly. This reading is

We will discuss Kripke's views on formal systems and his defense of the notion of 'self-evidence' in chapter 5. This is the paragraph mentioned by Kripke:

§48. 'Mathematical logic' has completely deformed the thinking of mathematicians and of philosophers, by setting up a superficial interpretation of the forms of our everyday language as an analysis of the structures of facts. Of course in this it has only continued to build on the Aristotelian logic. (*Remarks on the Foundations of Mathematics*, Part V)⁹³

The following passages from the *Remarks* may clarify what he has in mind:

§18. Now, what do we call 'inferences' in Russell or Euclid? Am I to say: the transitions from one proposition to the next one in the proof? But where is the *passage* to be found? – I say that in Russell one proposition follows from another if the one can be derived from the other according to the position of both in a proof and the appended signs – when we read the book. For reading this book is a game that has to be learnt. (*RFM*, Part I)

§19. One is often in the dark about what following and inferring really consists in; what kind of fact, and what kind of procedure, it is. The peculiar use of these verbs suggests to us that following is the existence of a connexion between propositions, which connexion we follow up when we infer. This comes out very instructively in Russell's account (*Principia Mathematica*). That a proposition $\vdash q$ follows from a proposition $\vdash p \supset q.p$ is here a fundamental law of logic:

$$\vdash p \supset q.p \supset .\vdash q$$

Now this, one says, justifies us in inferring $\vdash q$ from $\vdash p \supset q.p$. But what does 'inferring', the procedure that is now justified, consist in? Surely in this: that in some language-game we utter, write down (etc.), the one proposition as an

compatible with claiming that we follow rules blindly because we *say* that we follow rules but we don't have a rule *in mind* when performing the act itself. The *act* itself is blind.

⁹³ He also says: "[...] The formalization of logic did not work out satisfactorily. But what was the attempt made for at all? (What was it useful for?) Did not this need, and the idea that it must be capable of satisfaction, arise from a lack of clarity in another place?" (*RFM*, Part III, §85)

assertion after the other; and how can the fundamental law justify me in *this*? (RFM, Part I)

§20. Now Russell wants to say: “*This* is how I am going to infer, and it is *right*”. So he means to tell us how he means to infer: this is done by a *rule* of inference. How does it run? That this proposition implies that one? – Presumably that in the proofs in this book a proposition like this is to come after a proposition like this. – But it is supposed to be a fundamental law of logic that it is *correct* to infer in this way! – Then the fundamental law would have to run: “It is correct to infer ... from ...”; and this fundamental law should presumably be self-evident – in which case the rule itself will self-evidently be correct, or justified. “But after all this rule deals with sentences in a book, and that isn’t part of logic!” – Quite correct, the rule is really only a piece of information that in this book only *this* transition from one proposition to another will be used (as it were a piece of information in the index); for the correctness of the transition must be evident where it is made; and the expression of the ‘fundamental law of logic’ is then the *sequence of propositions* itself. (RFM, Part I)

And finally, a piece of advice from Wittgenstein:

§17. When we ask what inferring consists in, we hear it said e.g.: “If I have recognized the truth of the propositions ..., then I am justified in further writing down” – In what sense justified? Had I no right to write that down before? – “Those propositions convince me of the truth of this proposition.” But of course that is not what is in question either. – “The mind carries out the special activity of logical inference according to these laws.” That is certainly interesting and important; but then, is it true? Does the mind always infer according to *these* laws? And what does the special activity of inferring consist in? – This is why it is necessary to look and see how we carry out inferences in the practice of language; what kind of procedure in the language-game inferring is. For example: a regulation says “All who are taller than five foot six are to join the ... section.” A clerk reads out the men’s names and heights. Another allots them to such-and-such sections. – “N.N. five foot nine.” “So N.N. to the ... section.” That is inference. (RFM, Part I)

In the next three chapters we will discuss some of the problems that arise when we do *not* follow Wittgenstein’s advice and concentrate on the justification problem (our problem (a) in chapter 1) without pausing to “look and see how we carry out inferences in the practice of

language; what kind of procedure in the language-game inferring is.” Meanwhile, we buy some time to try to figure out whether we have something to say, not just about what inferring is not, but about what it could possibly be.

CHAPTER 4

Quine and Descartes on the Status of Logic:

Kripke's Critique

Kripke draws on the adoption problem to criticize three positions: Quine's views on the nature and justification of logic in "Two Dogmas;" Descartes's doctrine on the eternal truths; and Putnam's proposal to revise classical logic (which, for Kripke, presupposes a Quinean background and is, thus, subject to similar objections plus some others specific to his proposal).

In this chapter we will discuss Kripke's arguments against Quine (section 4.1) and Descartes (section 4.2). Both writers, especially Quine, take a stand on the nature of the logical truths and make a case for their justification. Kripke's use of the adoption problem to question Quine's and Descartes's understanding of the nature of logic will be our main focus, but we will also emphasize how these adoption-related criticisms ultimately weigh against their views on the justification of logic.

4.1.1 Quine's version of the Carroll Argument and the "Two Dogmas" Conception of Logic

We remarked before that interest in Carroll's note is by no means new or restricted to Kripke. The most famous application of Carroll's regress argument is in Quine's "Truth by Convention" (1936) (later also included in "Carnap and Logical Truth" (1954)). There, Quine directs his version of Carroll's argument against Carnap's linguistic doctrine of logical truth to great effect. Not only does he succeed at discrediting Carnap's specific proposal; the very idea that the truths

of logic could be based on some sort of stipulation was to become suspicious from then on. Kripke, nevertheless, thinks that the same objection that Quine raises against Carnap's proposal can be directed at Quine's own views in "Two Dogmas."

For Carnap, the *a prioricity* and necessity of logic, mathematics, and conceptual truths was grounded on linguistic conventions. He, together with the other members of the Vienna Circle, sought an explanation that afforded the distinction between sentences that were *a priori* and necessary⁹⁴ (which, for them, ran together) and the rest, without betraying their empiricist persuasion. Analyticity was the route chosen to accomplish this: mathematical, logical,⁹⁵ and conceptual truths were, unlike scientific statements, analytic, true in virtue of the meanings of their constituent expressions alone; and because they did not constitute truths about the world, they could be known *a priori*.

The route was hardly new. The idea of using the notion of analyticity as a way of dealing with the *a priori* can already be found in Leibniz's distinction between "truths of reason" and "truths of fact," or in Hume's distinction between "relations of ideas" and "matters of fact." The novelty was the element of linguistic conventionalism that Carnap and other members of the Circle introduced to account for the origin of analytic statements. For Carnap, the analytic statements themselves were linguistic conventions or the result of conventions, primarily selected in the light of pragmatic considerations.

The conventions were understood as deliberate stipulations (implicit definitions) of the truth of certain sentences, which were supposed to implicitly define expressions contained in them; the

⁹⁴ The idea was to avoid denying, as Mill had done before him, the existence of *a priori* knowledge.

⁹⁵ The members of the Vienna Circle were following the *Tractatus* in thinking that logical truths are tautological and, accordingly, do not express any facts.

stipulation of their truth was taken to constitute the meaning of some of those expressions.⁹⁶ By these means the *a priori* was intended to be smoothly incorporated into the empiricist framework: all *a priori* sentences were merely analytic, the result of our own linguistic conventions; as such they did not express any factual truths and their necessity could be explained on purely linguistic grounds.

What came to be known as Carnap's linguistic doctrine of logical truth consisted in the application of these ideas to the case of logic: the stipulation of a formal system's set of axioms and rules of inference as true is taken to implicitly define the logical constants of that system. There are, in Carnap's view, two different rules in the construction of an artificial language: formation rules, which specify what counts as a sentence of the artificial language, and transformation rules (or meaning postulates), which specify under what conditions an inference is legitimate. The latter are responsible for implicitly defining the logical constants of the system. Because their selection is purely conventional, the resulting linguistic framework reflects our own choices, which, as noted, are primarily guided by pragmatic considerations.

The analytic truths that constitute the linguistic framework are therefore only necessary and *a priori* justified in a framework-relative way. We are, Carnap argues, free to abandon a given framework and adopt another one. There is also an element of arbitrariness in the selection. Carnap defended a principle of tolerance according to which "it is not our business to set up

⁹⁶ Implicit definitions need not be stipulations that make the sentences containing the expressions to be defined true solely in virtue of their meaning. Boghossian (1996: 198ff) calls this conception of analyticity "metaphysical": an analytic statement "*owes its truth value completely to its meaning*, and not at all to the facts." (p. 198) In contrast, an alternative epistemological conception drops the idea that the truth of the sentence is established by stipulation but preserves a connection between the meaning of the sentence and its justification: our justification for holding them as true is given by its meaning alone. See also Boghossian (1997, 2003), Peacocke (1993; 2008: ch. 4), Hale and Wright (2000), Williamson (2007: ch. 3 and 4), and chapter 6 below.

prohibitions, but to arrive at conventions,”⁹⁷ and recommended “complete liberty with regard to the forms of language.” The decision to build or adopt a particular framework is therefore a contingent affair, mostly based on pragmatic considerations. But it might well also turn out to be arbitrary, given Carnap’s refusal to legislate what should be included and kept apart, or set a firm criterion for inclusion.⁹⁸

Now, towards the end of “Truth by Convention” Quine identifies a major difficulty for Carnap’s doctrine. If we follow Carnap in assuming that particular logical truths are established by means of conventional stipulations, Quine argues, we are faced with Lewis Carroll’s “vicious regress.” Quine’s point is that for those general conventions to be applied to particular cases, logical inferences will be needed in the metatheory, which in turn will *not* be the result of a stipulation. We could, at this point, introduce further stipulations, but the same problem would, of course, arise:

In the adoption of the very conventions [...] whereby logic itself is set up, however, a difficulty remains to be faced. Each of these conventions is general, announcing the truth of every one of an infinity of statements conforming to a certain description; derivation of the truth of any specific statement from the general convention thus requires a logical inference, and this involves us in an infinite regress. (Quine 1936: 96.)⁹⁹

And this ‘difficulty’ is later seen as insurmountable in “Carnap on Logical Truth”:

⁹⁷ He says:

... in logic there are no morals. Everyone is at liberty to build up his own logic, i.e., his own form of language, as he wishes. All that is required of him is that if he wishes to discuss it, he must state his methods clearly, and give syntactical rules instead of philosophical arguments. (Carnap 1937: 51-52.)

⁹⁸ For an exposition of Carnap’s views, see Coffa (1991), especially chapter 17. Coffa downplays Carnap’s commitment to the principle of tolerance.

⁹⁹ In “Truth by Convention” Quine also gives what he calls an “alternative” formulation of the argument. See fn. 130 below for discussion of this formulation.

... it is impossible in principle, even in an ideal state, to get even the most elementary part of logic exclusively by the explicit application of conventions stated in advance. The difficulty is the vicious regress familiar from Lewis Carroll which I have elaborated elsewhere. Briefly the point is that logical truths, being infinite in number, must be given by general conventions rather than singly; and logic is needed then to begin with in the meta-theory, in order to apply the general conventions to individual cases. (Quine 1954: 108)¹⁰⁰

Because there are infinitely many logical truths, they cannot be stipulated one by one. The conventions must be general. But then, if one wants to apply those conventions to individual statements, a logical inference connecting them is needed: we have to infer from the general convention stipulating the truth of a given schema that each of the instances of such schema is true.

Suppose we have a general convention stipulating that all replacements of the propositional variables ‘p’ and ‘q’ in ‘If p and q, then p’ by sentences are logical truths. How do we apply it to a particular case? Assume we have also the following statement, ‘If Alex is a lion and Gloria is a hippo, then Alex is a lion,’ and want to know if it is a logical truth. It would seem that some kind of logical inference connecting it with the general convention is needed. For example, an inference along the following lines: All replacements of ‘p’ and ‘q’ in ‘If p and q, then p’ by sentences are logical truths; ‘If Alex is a lion and Gloria is a hippo, then Alex is a lion’ has the

¹⁰⁰ In the foreword to Lewis (1969), Quine summarizes the argument as follows:

This is a characteristic and crucial case of appealing to convention where there can have been no thought of convening. For the philosophers in question count logical truth analytic; and here a circularity would arise if we were to take the conventions explicitly. The predicament is that in order to apply any explicit conventions that were general enough to afford all logical truth, we would already have to use logic in reasoning from the general conventions to the individual applications. (Quine 1969: p. xii.)

Note that he talks of “a circularity” here instead of an infinite regress.

form ‘If p and q , then p ;’ therefore, ‘If Alex is a lion and Gloria is a hippo, then Alex is a lion’ is a logical truth.¹⁰¹

The truths of logic were supposed to “spring” (Quine 1936: 96) from conventions, but the ‘difficulty’ is that we had to use logical inferences that were not given by conventions and hence remain unjustified. We could set up another convention to cover the logical inferences used before, but we will also need to make inferences in order to apply them, and so on. In Quine’s view, we can’t, not even “in principle,” account for the logical truths as springing out from explicit conventions in the manner suggested by Carnap. Perhaps, however, the conventions need not be explicit:

It may still be held that the conventions [...] are *observed* from the start, and that logic and mathematics thereby become conventional. It may be held that we can adopt conventions through behavior, without first announcing them in words; and that we can return and formulate our conventions formally afterward, if we choose, when a full language is at our disposal. (Quine 1936: 98.)

So the regress does not arise if conventions are directly adopted through behavior instead of having an explicit formulation. But Quine ultimately rejects this alternative because he thinks it is

¹⁰¹ This is obviously not the example considered by Quine. As Kripke pointed out to me, in this paper Quine has an unusual view of a formal system of logic – in this case, Lukasiewicz’s system for propositional logic. A formal system is understood as a specification of a class of statement of natural languages as true. Quine tries to derive statements of natural language such as “If time is money then time is money” from a set of conventions that correspond to Lukasiewicz’s postulates (which include some axiom schemes and *Modus Ponens* as a rule), and shows how this attempt leads us into an infinite regress. (Note that the system is complete in that all statements of natural language which are tautologous or logically true according to propositional logic can be obtained.)

The postulates are:

- (I) Let all results of putting statements for ‘ p ’, ‘ q ’, and ‘ r ’ in ‘If if p then q then if if q then r then if p then r ’ be true.
- (II) Let any expression be true which yields a truth when put for ‘ q ’ in the result of putting a truth for ‘ p ’ in ‘if p then q .’
- (III) Let all results of putting a statement for ‘ p ’ and a statement for ‘ q ,’ in “If p then if $\neg p$ then q ’ or ‘If if $\neg p$ then p then p ,’ be true.

unclear what the adoption of a convention prior to its explicit formulation as such would amount to, and, more importantly, because:

In dropping the attributes of deliberateness and explicitness from the notion of linguistic convention we risk depriving the latter of any explanatory force and reducing it to an idle label. We may wonder what one adds to the bare statement that the truths of logic and mathematics are a priori, or to the still barer behavioristic statement that they are firmly accepted, when he characterizes them as true by convention in such sense. (1936: 99)

The thought seems to be that conventions wouldn't be able to play the justificatory role Carnap intended. On this view, logical truths are not the result of deliberate linguistic conventions setting up an inextricable relation between the meaning of a sentence and its truth, but rather the result of regularities of behavior already in place and only later explicitly formulated. And it is hard to see, Quine seems to think, once the tight connection between the free stipulation of the meaning of a sentence and its truth (in the context of a framework) is lost, what could ground their truth. They would merely appear to be descriptions of what we do and we would have little reason to presume their truth.¹⁰²

¹⁰² David Lewis famously defended a view in which no explicit formulation for the adoption of a convention is required (Lewis 1969). In the Princeton Seminar Kripke mentions Lewis's proposal and says that he doesn't think it would work for logic (without spelling out the argument). One problem is that Lewis's account presupposes reasoning among those setting up conventions (see pp. 25 ff). For Lewis, a convention is a regularity in behavior, sustained by a system of preferences and expectations, that has a possible alternative. Conventions are set up by groups of people to solve "coordination problems" (such as driving on the right or left of a road). But figuring out other people expectations and desires, and alternatives to our behavior presupposes that such people are already in possession of a *logica utens*.

Note that Lewis rejects the idea of truth by convention: "say, if you like, that it is by convention that there are no *rightly so-called* married bachelors. But do not say that it is by convention that there are no married bachelors, in this world or any other. There couldn't be." (p. 207)

It is generally agreed that “Truth by Convention” has the rare privilege of actually managing to refute a philosophical view, in this case the linguistic doctrine of logical truth.¹⁰³ And Quine himself eventually came to believe that his objections to the Carnapian project called for very drastic adjustments that would lead to a final rupture with it.¹⁰⁴ The new model emerges in “Two Dogmas,” and the passage involving logic has been so widely quoted that I won’t attempt to resist the temptation to do so once again:

It becomes folly to seek a boundary between synthetic statements which hold contingently on experience, and analytic statements, which hold come what may. Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. Even a statement very close to the periphery can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws. Conversely, by the same token, no statement is immune to revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics; and what difference is there in principle between such a shift and the shift whereby Kepler superseded Ptolemy, or Einstein Newton, or Darwin Aristotle? (Quine 1951: 40)

The laws of logic are, just like our scientific hypotheses, open to revision. In principle every statement can be revised in the face of recalcitrant experiences. The main reason he gives for this comes a little earlier, when he writes:

The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only

¹⁰³ See for example Benacerraf (1973: 676) and Boghossian (1996: 217).

¹⁰⁴ According to Creath, “Truth by Convention” does not constitute “a radical break with analyticity or with Carnap’s program in general” (1987: 494). He argues that only later did Quine come to hold such views. However, Quine does express doubts in “Truth by Convention” on how the contrast between conventional and non-conventional truths should be understood. Moreover, as Creath himself would admit (1987: 498), it constitutes an important part of the background against which “Two Dogmas” was conceived.

along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery occasions readjustments in the interior of the field. Truth-values have to be redistributed over some of our statements. Re-evaluation of some statements entails re-evaluation of others, because of their logical interconnections - *the logical laws being in turn simply certain further statements of the system, certain further elements of the field*. (Quine 1951: 42-43. My emphasis.)

Scientific statements and logical and mathematical statements are viewed as making up a continuum. The law of universal instantiation, $2+2=4$, Newton's law of gravitation, and statements such as "Aconcagua is the tallest mountain in the Western hemisphere" and "there are brick houses on Elm Street" are all ultimately dependent on experience. In Quine's metaphor, they are all part of a web of belief, a structure with a periphery directly affected by empirical observations and an interior that becomes less affected by experience as we move deeper into it. But the elements of this continuum are not cut off from one another; there are logical connections among them. The revision of some statements may then entail the revision of others, and so the impact of experience, which is sharply felt at the periphery, may also be transmitted to the core of the system.

For Kripke, the key point is that Quine views the laws of logic as "*simply certain further statements of the system*." It is because they have no special epistemic status that they are open to revision, just like the rest of the system's statements. In contrast to Carnap, who, after the modern tradition, sought a principled way of making the distinction between *a priori* statements and the rest, Quine denies that there is a difference in kind among the statements in the web. Differences are just a matter of degree, of being more or less likely to be affected by experience depending – following Quine's figure – on how far from the periphery of the system a given statement happens to be.¹⁰⁵

¹⁰⁵ Hence, some of the statements mentioned above such as "there are brick houses on Elm Street" are closer to the periphery, and therefore more immediately responsive to observations. Some of them, like

The statements' epistemological homogeneity is needed for Quine's holism: it is because there is no seam in the web of belief (Quine and Ullian 1970) that the logical and mathematical principles can be thought to face the tribunal of experience together with the natural laws, auxiliary hypothesis, etc., and so be confirmed when the total scientific theory is empirically corroborated. If there is no sharp distinction between logical and mathematical statements on the one hand, and scientific statements on the other, there is no reason to suppose that the former need to be justified in a special way.¹⁰⁶ As in the case of highly theoretical scientific statements, the empirical evidence that counts in their favor is always indirect. The decision to include certain logical principles and mathematical statements as part of the system is not essentially different from the decision to include scientific statements. It is made on the basis of the simplicity and explanatory power they bring to the system. And it is also the empirical success of the system as a whole that ultimately justifies them.¹⁰⁷

As Devitt (2011b) is keen on pointing out, the model leaves "no place" for *a priori* justification. Following Carnap, Quine identifies *aprioricity* with analyticity, and defines analytic statements as

"2+2=4" or the law of universal instantiation, are located deep within the interior of the structure and so less likely to be defied by observations. Between such two extremes, there is a "grading off" (Quine 1960: 42-44) of the statements' vulnerability to observations, but since there are logical connections among them, the whole web of belief is ultimately shaped in response to experience.

Dummett (1973c: 376-377) argues, however, that the distinction between center and periphery vanishes due to Quine's confirmation holism and his claim that any statement can be either revised or preserved in the face of recalcitrant experience. See also fn. 122 below.

¹⁰⁶ But Quine's holism also lends support to his epistemic egalitarianism towards statements by giving a unified answer to the question of how the statements are to be justified. It also provides a way of avoiding the implausible claim, usually associated with Mill, that mathematical statements and logical principles are established on the basis of direct inductive generalizations.

¹⁰⁷ Quine's epistemic holism was anticipated by Duhem, as he himself notes (see Quine 1951: 41). In *Word and Object*, though, he concedes that there was an "excessive holism espoused in occasional brief passages of mine" (Quine 1960: 13, fn. 5). In a later writing he says that "looking back on it ["Two Dogmas"], one thing I regret is my needlessly strong statement of holism" (Quine 1991: 268), and talks about a moderate holism in which a statement does not need the whole of science to be corroborated but "chunks" of it, clusters of statements that "imply an observable effect of an observable experimental condition." (Quine 1991: 268)

“a limiting kind of statement which is vacuously confirmed, *ipso facto*, come what may.” (1951: 41)¹⁰⁸ But there is no application for this concept in the model: any statement, even if empirical observations appear to clash with it, can be preserved. We can always choose to revise other statements to accommodate the problematic empirical observations. And that’s only one side: any statement may be revised or abandoned as well in an attempt to improve the empirical success of the system as a whole.

Of course Quine makes it clear that since a revision of the logical laws or mathematical statements would require more extensive adjustments in our body of beliefs than other revisions, they should only be undertaken as a last resort. His “maxim of minimum mutilation” recommends that those beliefs which are least crucial to our overall body of beliefs should be rescinded first.¹⁰⁹ Simplicity and conservativeness are the guiding principles of belief revision, the “considerations of equilibrium” that should be taken into account, and can usually be expected to secure the mathematical and logical statements of the system. Yet the maxim is simply a methodological constraint, which can ultimately be disobeyed. After all, “no statement is

¹⁰⁸ In Quine’s view, *a priori*ity, analyticity, and incorrigibility go together. I have already claimed (fn. 26) that, to my mind, it is a mistake to identify *aprioricity* and incorrigibility.

¹⁰⁹ See Quine (1970b: 7, 86, 100); and Quine (1960: 20). He writes: “... suppose that from a combined dozen of our theoretical beliefs a scientist derives a prediction in molecular biology, and the prediction fails. He is apt to scrutinize for possible revision only the half dozen beliefs that belonged to molecular biology rather than tamper with the more general half dozen having to do with logic and arithmetic and the gross behavior of bodies. *This is a reasonable strategy* – a maxim of minimum mutilation. But an effect of it is that the portion of theory to which the discovered failure of prediction is relevant seems narrower than it otherwise might.” (Quine 1970b: 7)

Gary Ostertag has suggested to me that the maxim is purely descriptive of our practices. I disagree. Quine does indeed talk of “our natural tendency to disturb the total system as little as possible” (1951: 43), but that does not deprive the maxim of its intended normativity. There is no incompatibility in claiming that we tend to observe the maxim and should observe it at the same time. Calling it “maxim” and saying that it is “a reasonable strategy” (in the quote above) suggests exactly that. Whether it affords a descriptive reading or a normative one won’t affect the arguments given below (in fact I think that, if anything, the former reading would make the objections worse).

immune to revision.”¹¹⁰ If empirical findings suggest a revision of our most entrenched logical and mathematical statements, a revision that augments the explanatory power of the overall system and which could not be achieved through other more minimal possible changes, then the maxim could in principle be ignored.¹¹¹ But according to Quine, it is a rather intuitive maxim, and it explains the “air of necessity that attaches to logic and mathematical truth.” (Quine 1970: 100) It is our determination to observe the maxim what accounts for their apparent necessity.¹¹²

¹¹⁰ Quine’s position is more complicated because his *Philosophy of Logic* (Quine 1970b: ch. 6) he appears to contradict the “Two Dogmas” pro-revisionist attitude, claiming that a change of logic is a change of subject. He denies, however, any tension: “at this level a change of theory is *itself* a change of meaning, though not always conversely. If we just write ‘and’ for ‘or’ and vice versa, we change meanings but not theory. If we abandon the law of excluded middle, we change meanings *and* theory, the law does not survive in any rendering.” (Quine 1995: 352) He makes similar remarks in Quine (1954: 113). Still, Quine’s change of subject argument makes hard to see how a rational revision of the laws of logic could actually occur. For Quine, in order to interpret and understand (translate, in his sense) we must proceed in accordance with the maxim ‘save the obvious.’ Given that the laws of logic are obvious (see fn. 113 below) and that it is a central part of the translation practice to preserve obvious truths, they must be preserved by any correct translation. The problem is that this view seems to make rational disagreement over a logical law impossible. As Levin remarks, “the upshot of the translation argument appears to be that the logical laws cannot be denied, that we *cannot* coherently describe the situation the empirical view so equably envisions.” (1978: 47. In the rest of the article Levin discusses how this tension can be resolved. See also Stroud 1968).

¹¹¹ Dummett argues that Quine is not denying that philosophical criticism could ever lead us to revise the forms of reasoning accepted in classical logic: “no empiricist, however dogmatic, will hardly be prepared to deny in advance than any analysis could ever have that upshot.” (Dummett 1976: 269) According to Dummett, Quine’s focuses on revision as a response to experience instead of revision based on ‘philosophical criticism,’ because that was the kind of revision resisted by the empiricists he was criticizing. But this interpretation seems to reflect Dummett’s own agenda more than Quine’s. Because the web of belief is conceived as being shaped in response to experience, it hard to see how anything else but empirical considerations could have the last word when deciding which logical principles are ultimately kept in the system.

One of the problems is that the question of *how exactly* empirical considerations would lead us to revise logical principles is left unclear. In the quoted paragraph, Quine mentions the proposal to revise the law of excluded middle as a means to simplifying quantum mechanics. But how the maxim of minimum mutilation and considerations of simplicity could in this case be outweighed for the benefit of the system as a whole is left without explanation. It is fairly clear from later writings that Quine did not view this case as a serious contender for revision (see, for example, Quine 1970b: 86). For a summary of some problems related to the maxim, see Sober (2000: 262-263).

¹¹² Quine also holds that “every logical truth is obvious, actually or potentially. Each, that is to say, is either obvious as it stands or can be reached from obvious truths by a sequence of individually obvious

4.1.2 The Standard Objection

A common criticism of Quine's attempt to give an empirical justification of the logical truths is that inferential transitions will have to be presupposed in setting up the web and making it responsive to observations. Justification requires the confirmation of the system as a whole (but see fn. 108 above). Yet, if in setting the system up and making it responsive to observations, we have to rely on inferences that correspond to those very logical truths, the empirical justification of at least some of those truths will appear to be circular.

Call this criticism the standard objection. It has been raised in different ways, mostly in connection to the possibility of carrying out a revision of the logical truths in Quine's system.¹¹³ Katz, for example, argues that Quine's model of belief revision presupposes the principle of non-contradiction, which "mandates revision when there is inconsistency." (2000: 72) Non-contradiction is in his view the rationale for belief revision and what indicates *when* "the system as a whole must be restored to consistency within itself and with its observational periphery;" (p.72) it is thus needed "to initiate the process of revising presently accepted statements, otherwise we have to tolerate a radically laissez-faire epistemology on which anything goes." (p.

steps." (Quine 1970b: 82-83) Since for him the logical truths don't have a special epistemic status, 'obvious' is not understood in an epistemological sense, but used in an "ordinary behavioral sense" (p. 82) to mean what pretty well everyone in the cultural community unhesitatingly assents to. According to Quine, the two considerations that can be given in support of Carnap's "linguistic doctrine" – "that alternative logics are inseparable practically from mere change in usage of logical words" and "that illogical cultures are indistinguishable from ill-translated ones" (1954: 105) – are explained by the obviousness of the logical truths.

¹¹³ See, for example, Katz (2000), Wright (1986), Boghossian (2000), Field (1998, 2000), Bonjour (2001), Sober (2000), and Shapiro (2000).

73) Because of this central role, however, the principle is subject to ‘the revisability paradox’:¹¹⁴ given that any argument for belief revision presupposes the principle of non-contradiction, if the principle itself were indeed revisable, the argument for its revisability would be unsound; it would rest on a false principle. Non-contradiction is therefore not revisable because no argument for its revisability can be sound. Yet, since according to Quine any belief is revisable, the principle must be revisable, and “hence it is both revisable and not revisable.”¹¹⁵ (p. 74)

Boghossian,¹¹⁶ on the other hand, has argued that in Quine’s account “a large number of the core principles of logic will have to be used to select the logic that [...] is maximally justified by experience.” (Boghossian 2000: 233) The selection is supposed to occur in the context of an evaluation of our overall theory and on the basis of considerations such as simplicity, conservativeness, and predictive power. But the evaluation is meant to be internal: the theory’s underlying logic is used to draw the theory’s consequences, which are ultimately contrasted with the observation sentences. Now, suppose, Boghossian says, that we have a theory *T*, a logic *L*, and consequence *p*. A recalcitrant experience, however, leads us to believe not-*p*. In trying to fix this up we will have various options, where changes of *T*, *L*, or both, are possible. So we will have to select among various ordered pairs such as $\langle T', L \rangle$, $\langle T, L' \rangle$, $\langle T'', L \rangle$, etc., the pair that best accommodates the considerations of simplicity, conservativeness, predictive power, and

¹¹⁴ According to Katz, Quine’s epistemology of belief revision is based on three principles that indicate when, where, and how a revision should proceed. The when-principle is the principle of non-contradiction, the where-principle is the principle of universal revisability (which states that no statement is immune for revision), and the how-principle is the principle of simplicity. The three principles are subject to “the revisability paradox”: “Since the constitutive principles are premises of every argument for belief revision, it is impossible for an argument for belief revision to revise any of them because revising any one of them saws off the limb on which the argument rests. Any argument for changing the truth value of one of the constitutive principles must have a conclusion that contradicts a premise of the argument, and hence must be an unsound argument for revising the constitutive principle.” (p. 73) See Adler (2003) and Colyvan (2006) for replies to Katz.

¹¹⁵ Katz also offers a version of the paradox for the case in which the principle is not rejected as false but simply dropped (see p. 74).

¹¹⁶ See Boghossian (2000: 233-234). His argument follows Wright (1986: 192-193).

so on. The difficulty Boghossian points out is that selecting the winning pair will require inferences taking place in the meta-theory. UI and MPP inferences are, for example, good candidates for being unavoidable.¹¹⁷ But then, Boghossian notes, if circularity is to be avoided, our belief in *at least* some basic logical principles¹¹⁸ is not explained by reference to their empirical success, since inferences that accord with those principles were instrumental in establishing the winning theory and its underlying logic. Moreover, an additional problem stressed by Field (1998: 12) is that the logic used in the meta-theory will tend to favor a theory with that underlying logic.

Though the standard objection is usually raised in connection to the revision of the logical principles – and therefore as a way of challenging Quine’s case against the *aprioricity* of logic – it can also be seen as an objection to their empirical justification in Quine’s model of confirmation.¹¹⁹ For it seems that even if there are no competing theories with different underlying logics, we still need to determine whether, for example, the consequences drawn from the theory using its own logic fit the observation sentences or not, or whether a reasonable degree of simplicity and predictive power has been achieved. And it is hard to imagine that this could be done without relying on background UI or MPP inferences. Once again, the choice seems to be between giving up an empirical justification for all the logical truths, or accepting its

¹¹⁷ Boghossian claims that:

... instances of the following forms of reasoning are presumably unavoidable in the meta-theory:
 The best set of observation sentences is the set with property F. Set O has property F. Therefore, O is best.
 <T*, L*> is that theory-logic pair that predicts O. O is the best set of observation sentences.
 Therefore, <T*, L*> predicts the best set of observation sentences. (2000: 233)

¹¹⁸ Boghossian emphasizes that the argument does not necessarily work for all principles, since not all the corresponding inferences will be needed in the evaluation of the different theories and their underlying logic. But he claims that it works for a “sizeable number of the core principles.” (2000: 234)

¹¹⁹ For a general criticism of Quine’s views on confirmation, see Sober (2000: 263ff).

circularity. Yet, for Quine the latter choice would amount to a recognition that his own proposal does not fare better than Carnap's – at least when it comes to the main “difficulty” he himself identifies in the “linguistic doctrine of logical truths.”

Two problematic aspects of Quine's model, both of them repeatedly pointed out, may obscure the point raised by the standard objection. One of them is the vague and metaphorical nature of Quine's account of the model.¹²⁰ We already pointed out (fn. 112) that he does not explain how exactly revisions are supposed to be carried out. Rey notes (1998: 30) that Quine himself admitted that the five “virtues of hypothesis,” that is, simplicity, generality, modesty, conservatism, and falsifiability, are described in a rather imprecise way (Quine 1986: 493).

The problem with this is that, because the methodological considerations and procedures we are supposed to follow when evaluating the system are delineated in a highly imprecise way, it is hard to say if attempts to show that the use of some logical principles will be presupposed do indeed give a correct characterization of the inner workings of the system. Yet merely rejecting such characterizations as a proper account of how the system is supposed to operate doesn't seem to be enough. The burden of proof appears to be on the Quinean camp: unless an argument explaining *how* an empirical justification of *all* the logical truths in the context of the web could even be possible, we seem to have little reason for endorsing his empirical account or hoping that it will turn to be correct.

¹²⁰ See, for example, Field (1998: 13ff, 2000: 129) and Dummett (1973c: 375-376). Rey says: “Traditional accounts of knowledge of logic and mathematics have indeed been obscure. But no one should be under any illusion that Quine's account is really any better. At best, it avoids obscurity only by so distancing the issues that all the crucial details are lost.” (1998: 31) Even Devitt admits that “the vagueness cannot be denied.” (1998: 61) But he claims: “it is hard to see why this vagueness constitutes an effective criticism of the Quinean view of logic. Perhaps it would be if realistic alternative evidential systems that treated logic as *a priori* had been spelled out.” This doesn't seem very fair considering that empiricists point to the obscurity of the accounts of the *a priori* based on rational intuition as the main reason for rejecting them. If both accounts are obscure when it comes to the specific case of logic, then obscurity cannot be used as a decisive argument by either side.

The second aspect, also noted by many,¹²¹ is that Quine does not draw a sharp distinction between the statements of the logical truths and the inferential transitions that connect the rest of the statements of the system. Given Quine's criticism of Carnap, his failure to stress this distinction is certainly puzzling. But the relevant passages don't leave space for doubt. Consider:

Re-evaluation of some statements entails re-evaluation of others, because of their logical interconnections – the logical laws being in turn simply certain further statements of the system, certain further elements of the field. Having re-evaluated one statement we must re-evaluate some others, which may be statements logically connected with the first or may be *the statements of logical connections themselves*. (1951: 42. My emphasis.)

¹²¹ See, for example, Priest (1978: 293) and Shapiro (2000: 336-337). Dummett (1973d: 596) also appears to have noted this point when he argues that the "Two Dogmas" model cannot account for the inferential connections between the statements in the web. He writes:

Quine's thesis involves that the principles governing deductive connections themselves form part of the total theory which, as a whole, confronts experience. Presumably, in order to avoid Achilles and the Tortoise troubles of the Lewis Carroll type, we must recognise the total theory as comprising rules of inference as well as logical laws in the form of valid schemata and their instances: but there is no reason to suppose that Quine draws a distinction between the status of such rules as against laws like Excluded Middle; they too must be equally liable to rejection under a heavy enough impact from without. But, in that case, there is nothing for the inferential links between sentences to consist in. They cannot be replaced by superinferential links, compelling us, if we accept certain logical principles, to accept also the consequences under those principles of other sentences we accept: for any such superlogical laws could in turn be formulated and considered as sentences no more immune to revision than any other. (1973d: 596)

[...] As a result of the second thesis [that no statement is immune from revision], the internal structure of the theory, consisting in the interconnections of sentences with one another, is totally dissolved, and the theory becomes a mere featureless collection of sentences standing in no special relations to each other. (1973d: 597)

And in Dummett (1973c) he claims that:

For inferential connections of a deductive character, at any rate, we know of no way of formulating the existence of such connections that will not have, at least as a by-product, the effect of guaranteeing the truth of certain sentences, which will be precisely the analytic ones. Hence, a thesis that denies the existence of analytic sentences calls in question the existence of any inferential links between non-analytic sentences. (1973c: 377)

or,

Even a statement very close to the periphery can be held true in the face of recalcitrant experience by pleading hallucination or *by amending certain statements* of the kind called logical laws. (1951: 43. My emphasis)

And even more explicitly, in *Word and Object*, he says:

In an obvious way this structure of interconnected sentences is a single connected fabric including all sciences, and indeed everything we ever say about the world; for the logical truths, at least, and no doubt many more commonplace sentences too, are germane to all topics and thus *provide connections*. (Quine 1960: 12-13. My emphasis.)¹²²

These passages not only make clear that Quine neglects the distinction between the statements of the logical truths and the inferential transitions, but also that he thinks the key role is played by the statements and only grants a derivative status to the inferences themselves. In his view, the *logica docens* generates the *logica utens*; the logical truths “provide” the connections. This point is relevant to the standard objection because it makes the issue of logical revisionism appear simpler than what it might otherwise be. On this view, revising “the statements of the logical connections” has an immediate effect on the inferential connections themselves. Hence, when facing a recalcitrant experience, it would seem that we could just go to the core of the system, modify a logical truth, and by doing so change all the inferential connections. Once the distinction is clearly drawn, however, it becomes harder to see how a revision of all inferential patterns could be carried out.

¹²² One can only wonder what these “many more commonplace sentences” providing connections, like the logical laws, could possibly be. It is hard not to suspect that they are thrown in to avoid ascribing the logical laws the rather special and exclusive role of providing the connections of the system. That alone would set them apart.

A die-hard Quinean is unlikely to be moved by the standard objection or the two problems we have pointed out. Regarding Quine's neglect of the distinction between logical truths and the inferential transitions, he may want to claim that once it is clearly introduced, we will still be able to revise (or confirm, for that matter) the inferences themselves. But there are reasons to suspect that the issue won't be so simple (see below). He is also unlikely to be seriously troubled by the vagueness involved in the description of the web and its inner workings (see fn. 121). And about the standard objection itself, he may want to follow Quine and claim that confirmation (or revision) doesn't happen all at once, but bit by bit.¹²³

References to Neurath's boat are symptomatic of this attempt to overcome the standard objection: in the same way that we can rebuild a boat plank by plank by staying afloat on it, we can confirm (or revise) the different parts of the web of belief by 'chunks' (see fn. 108), clusters of statements large enough to "*imply* an observable effect of an observable experimental condition." (Quine 1991: 268. My emphasis.) The following passage from Devitt gives an idea of how this plank by plank strategy is supposed to work:

The Quinean solution to this problem is captured by the famous image from Neurath of rebuilding a boat whilst staying afloat on it. We can rebuild any part of the boat but in so doing we must take a stand on the rest for the time being. So we cannot rebuild it all at once. Analogously, we can contemplate revising any part of our evidential system S but in so doing we must hold fast to the rest of the system for the time being. So we cannot contemplate revising it all at once. This applies to the logic that is built into S. Field rightly insists that we need some such logic to draw any conclusions from the empirical evidence. Still, we could contemplate revising T and the system S that it recommends. Suppose that system S', recommended by F, holds fast to most of classical logic but weakens the distributive law. We could employ S' to test our total science against the empirical

¹²³ Another possibility is to defend a circular justification (see Devitt 2011b: 281ff). We will discuss this possibility in some detail in chapter 6 (in the context of Boghossian's concept – constituting account). But it seems to me that it is rather doubtful that a circular justification would be a good match for the Quinean. See fn. 188 below.

evidence. If total science comes out better on this test than our present test employing S , then we have empirical grounds, of the usual Duhem-Quine holistic sort, for preferring S' . (Devitt 1998: 60-61)

In a later paper, however, Devitt acknowledges that “in thinking about this it is important to remember that S must include rules governing its own potential replacement, rules governing the choice between TS and its rival TS' that recommends a different system S' . It is hard to see how these rules, vaguely indicated by the Neurath metaphor, could themselves be justified or revised in the Neurath way.” (2011b: 281)

So even some *very* die-hard Quineans are willing to admit that adding a metaphor to the already obscure picture introduced in “Two Dogmas” doesn’t go a long way in helping us see how the logical principles and the corresponding inferences used in the confirmation of a theory can themselves be empirically justified. What we need is an argument showing that this is possible *in the specific case of logic*.¹²⁴

¹²⁴ Devitt (2005: 110) offers the following reply to what he calls the “self-defeat problem”:

... there are reasons for thinking that we can indeed show an evidential system to be defective using that very system. First, it seems undeniable that our evidential systems have changed. (i) A good deal of the impressive scientific progress over the past three centuries has been in improved methodologies: we have learnt a vast amount not only about the world but also about how to learn about the world. As a result, much education of the young scientist is in these methodologies: think of physics and psychology, for example. (ii) Educated folk have tried to adjust their thinking in light of evidence that we normally tend toward certain sorts of irrationality; for example, counter-induction, and ignoring base rates in thinking about probabilities. (iii) Even our deductive practices have been affected by the rise of modern logic. Next, the process of making any of these system changes must have been governed by some evidential system, the one that was then current. So, that system was used to establish an epistemological thesis that led to the system’s replacement. These examples give us good reason to think that an evidential system could be used rationally to undermine itself. Accepting the non-epistemological part of our web and governed by S as usual, we find T wanting and so replace it and the system S that it recommends.

Devitt’s argument would be more persuasive if we assume that our grounds for basic logical principles and the corresponding rules of inference are no different than the grounds for other rules or beliefs in our “evidential system.” But that’s exactly the point in dispute. This is one of the reasons why we need an

4.1.3 “Two Dogmas” and the Adoption Problem

The standard objection is not, in any case, the only – or, in my view, the most important – objection to Quine’s views on the status and justification of logic. But it should already suggest, it seems to me, that we cannot think of the basic principles of logic as being in the same league with the other statements of the system. And this is exactly Kripke’s main quarrel with the “Two Dogmas” account of logic.

Kripke begins his discussion of Quine’s views on logic by remarking that the idea that we have an open choice concerning the revision of the statements in the system is very strange.¹²⁵ He quotes a passage from Gardner (1971) to illustrate the point. Talking about Quine’s maxim of minimum mutilation, Gardner says:

The maxim is that when we have to choose among alternative revisions in a theory to account for unexpected observations, similarity of the new theory to the old counts heavily in its favor. One way to avoid unnecessary radical revisions is to abandon the less general beliefs used in deriving the false prediction. If we find a white raven, we abandon ‘All ravens are black’ rather than quantification theory, which applies to our entire world-theory and ought not to be held responsible for

argument for the specific case of logic. Moreover, even if “our deductive practices have been affected by the rise of modern logic,” this does not show that whatever modifications may have taken place are the result of empirical considerations. Conceptual changes – as the one discussed in the next chapter concerning Aristotelian logic – need not be problematic for the *a priorists*.

¹²⁵ Kripke’s reconstruction of Quine’s views on the status of logic goes as follows:

There is no sharp cleavage between logic and other domains of scientific inquiry or what have you. They are all of a piece. Logic is just one of a variety of tools we use in organizing our experience. So far, perhaps, classical logic, which is the logic we have originally adopted, has well stood the test of experience. No doubt it should not be abandoned for trivial reasons. Maybe it has been well confirmed, since by using classical logic we have been able to make a large number of predictions which are well justified. We *could*, however, in principle and in response to recalcitrant experience, choose to revise our logic rather than to revise something else, like physics or geometry, and perhaps some day experience will lead us to do so. (Princeton Seminar)

this failing of the biological part of it. (1971: 526)

Once granted that this raven is white, we can either give up the statement “all ravens are black” or abandon the *statement* of the law of universal instantiation. The methodological maxim recommends one option over the other, but we may choose to disregard it. The point is that the choice is open to us: we may well decide to stick to “all ravens are black” and give up the *statement* of UI. We have seen that Quine is quite explicit about this: “Having re-evaluated one statement we must re-evaluate some others, which may be statements logically connected with the first or may be *the statements of logical connections themselves*.” (1951: 42. My emphasis.)

At least the following choices would seem to be consistent with Quine’s model: (1) if we find a white raven we may decide to ‘re-evaluate’ the statement “all ravens are black” because it is ‘connected’ with “this raven is black,” which contradicts the now true statement “this raven is white.” And for all we know we may also have to re-evaluate biological statements on ravens. (2) Perhaps, however, we are not in the mood for doing “re-evaluations” and decide to reject the statement “This raven is white” by, say, pleading hallucination, defective vision, or simply painting the raven black. (3) Another possibility is to re-evaluate the statement of UI. We may be fond of “all ravens are black” because, say, we have been developing a biological theory with wide theoretical implications that explains why ravens can only be black.

For option (3) to work, abandoning the *statement* of UI would have to block the *inference* from “All ravens are black” to “This raven is black.” And Quine seems to think that this is easily done by re-evaluating “the statements of the logical connections themselves” – in this case, the statement of UI. But then it is only if we are committed to the UI principle that “This raven is black” follows from “All ravens are black;” otherwise, how would my dropping the principle

affect the logical connections themselves? (Note that this point is related to the problem we stressed before: Quine's failure to draw a sharp distinction between the logical truths and the inferential transitions.) About this, Kripke says:

Just because we believe all crows are black, that doesn't in itself commit us to believing that this crow is black. It's only if we believe that all crows are black *plus* universal instantiation, that we are committed to believing that this particular crow is black. All we have to do to reject this conclusion is to deny, or doubt, or at least hold in suspended judgment, the law of universal instantiation; and then it will be doubtful whether this conclusion really follows, and we will certainly not be committed to it.

[...] if we just believe all crows are black, we are not *ipso facto* committed to concluding that this crow is black. We have a choice: to either go and deduce this or revise our logic so that this conclusion doesn't follow. This means that, in the absence of this particular statement in the system of interconnecting statements, one would not be able to conclude this crow is black.

Kripke remarks that since for Quine the logical principles are subject to the same sort of empirical considerations as the rest of the statements, yielding a correct prediction not only corroborates the relevant scientific hypothesis but also the logical principles used in the inferences connecting the hypothesis to the prediction. Logical principles are, however, deeply entrenched in the system, given that they have received confirmation in varied and multiple areas of the totality of science.¹²⁶ It is because in addition to accepting the principle as part of the

¹²⁶ Kripke quotes a passage from Berry (1969) to emphasize this point. Berry says:

... the distinction between the *a priori* (or analytic) and the *a posteriori* (or synthetic) cannot in general be drawn. What has passed for the one differs from what has passed for the other only in being less affected by observation, but this is merely a matter of degree. The logical rule of universal instantiation, the Pythagorean theorem and Newton's law of gravitation display in ascending order varying degrees of dependence upon the world of sensory data. Yet these three are similar in that, although none confronts that world directly, each mediates between the so-called observation sentences that do; witness the fact that each of the three principles may play its role in the single prediction of one observed position from others. Each such principle provides connective tissue invaluable in passing by inference or computation from observation to

system we also use it that “one can say that this law is being constantly applied and constantly confirmed, or being shown to be fruitful in the empirical predictions that it is used to make.”

(Princeton Seminar)

In Kripke’s view, the Quinean model assumes that either at some point (probably very early on) we adopted the principles of (classical) logic, or that perhaps they are inborn in us. The important point is that according to this view, it is *because* we chose to accept them as part of our web of belief that we are committed to them and their consequences. Yet our commitment is essentially similar to our commitment to well established scientific hypotheses and, hence, provisional: we could choose to drop them or adopt other principles and then we would no longer be committed to those consequences.¹²⁷ [Notice here that, in terms of the adoption problem, it would be irrelevant to insist that our commitment arises out of its being true or well confirmed, the question is whether the acceptance of a logical law can by itself generate such commitments.]

observation. Each is justified pragmatically for each is tied up with observation by being part of the overarching theory which is science as a whole, itself empirical because at its edges (to use Quine’s expression) it *is* immediately responsive to observation. If we adopt Quine’s holistic view of science, it seems that we can argue that there is nothing a priori about logic and that consequently we do not need to adopt a position which explains its apriority (1969: p. 246)

Kripke notes that, according to Berry “the rules of logic, the rules of geometry, and the rules of physics are all used in making predictions. We use logic in making deductions from hypotheses.”

¹²⁷ Consider, for example, the following passage by Putnam (for more on Putnam’s views and Kripke’s criticisms, see chapter 5):

... instead of arguing: ‘classical logic *must* be right; so something is wrong with these features of quantum mechanics’ (i.e. with complementarity and superposition of states), one may perfectly decide ‘quantum mechanics may not be right in all details; but complementarity and superposition of states are probably right. If these are right, and classical logic is also right, then either there are hidden variables, or there is a mysterious cut between the observer and the system, or something of that kind. But I think it is *more likely that classical logic is wrong* than that there are either hidden variables, or “cuts between the observer and the system”, etc.’ (Putnam 1968: 189)

Already guessing where Kripke is going with all this, a die-hard Quinean is likely to complain that Quine never said that we adopted the laws of logic. But the idea that we can block an inferential connection between the statements in the system by dropping or changing a logical law indicates that it is because we accept the principles that the corresponding inferential transitions are taken to hold. Quine's failure to stress the distinction between the inferential connections and the laws of logic, and his claim that the latter "provide" the connections also reinforce this interpretation.

Moreover, though Quine rejects the idea that we can stipulate the *truth* of the logical principles by means of conventions, he retains a role for them as a way of introducing statements in his web. In "Carnap and Logical Truth" he introduces a distinction between legislative and discursive postulations. A legislative postulation,

... institutes truth by convention, and seems plausibly illustrated in contemporary set theory. On the other hand discursive postulation is mere selection, from a preexisting body of truths, of certain ones for use as a basis from which to derive others, initially known or unknown. (p. 111)

Later on he says that deviations from the law of the excluded middle,

... insofar as meant for serious use and not just as uninterpreted formal systems, are as clear cases of legislative postulation as the ones in set theory. (p. 113)

But the fact that they are the product of a postulation does not make them true by convention. The conventional element concerns the *act* of postulating them, not the *grounds* for their truth. Quine asks:

... might we not still project a derivative trait upon the sentences themselves, thus speaking of a sentence as forever true by convention if its first adoption as true was a convention? No; this, if done seriously, involves us in the most unrewarding historical conjecture. Legislative postulation contributes truths which become integral to the corpus of truths; the artificiality of their origin does not linger as a localized quality, but suffuses the corpus. (p. 112-113)

And then he says:

...we have found some sense in the notion of truth by convention, but only as attaching to a *process of adoption*, viz., legislative postulation, and not as a significant lingering trait of the postulated sentence. [...] we note legislative postulation in set theory and, at a more elementary level, in connection with the law of excluded middle. (p. 114. My emphasis.)

But this in no way distinguishes logic or set theory from the natural sciences, since the theoretical hypotheses of natural sciences can also be introduced by legislative postulations,

... for surely the justification of any theoretical hypothesis can, at the time of hypothesis, consist in no more than the elegance or convenience which the hypothesis brings to the containing body of laws and data. (p. 114)

Though for Quine our acceptance of the logical truths is a consequence of their obviousness, he claims that,

... a day could likewise have been when our elementary logic was itself instituted as a deliberately conventional deviation from something earlier, instead of evolving, as it did, mainly by unplanned shifts of form and emphasis coupled with casual novelties of notation. (p. 113)

So Quine doesn't seem to think that there is a problem with the idea that logical truths – and therefore, in his view, the inferential transitions as well – could originate in conventions. But once

in the web, they are on probation (though doing fairly well so far, at least in Quine's opinion), as is any other hypothesis to be found there. Whether we retain them or not is a matter of empirical fertility, simplicity, and so on.

No one will fail to suspect at this point that Kripke sees things a little differently. He argues that if basic logical principles are understood as hypotheses or beliefs that we adopt in order to (indirectly) determine their empirical impact, we should conclude that their scientific fertility is zero. They will not lead to a single prediction:

I want on the contrary to hold that, so regarded, the law of universal instantiation is completely useless. It has never led to a single prediction, or been of any use to us whatsoever, and so it cannot be said to be confirmed. Moreover, we cannot be thought of as having adopted it. If we did adopt it, it would have done us absolutely no good. (Princeton Seminar)

[...]

Now what does this show? Does this show some sweeping conclusion, like "Logic is not revisable"? I certainly don't want to lead to some big conclusion like that. But it *does* show that if you start talking about logic on the analogy with just anything else in science, that's not so easy a thing to do. It has a very different character because, without being able to reason, we can't say what are the consequences of a hypothesis, or what will happen if we adopt this rule or that. You have to reason in order to say what would be the case *then*. And there are certain rules which you just couldn't adopt: you couldn't tell them to yourself, because if you told them to yourself without already using them, they would be useless; so they either don't help you or they were superfluous anyway. Universal instantiation has this character. If anyone tells it to you, it is superfluous –that is, superfluous for getting you to perform any particular inference.

As in Harry's case, we would already need to infer in accordance with them to find the logical beliefs or hypotheses of any use. Otherwise, they would be worthless, inert, and as such they

wouldn't ever turn out to be confirmed or disconfirmed simply because they won't lead us to any conclusion. And, *pace* Quine, it is this aspect what sets them apart from all the other statements of the web and makes them unique.

Now, not content with arguing that the logical truths don't "provide" any connections, Kripke goes on to argue that Quine should have been aware of this criticism:

... the Carnapian tradition about logic maintained that one can adopt any kind of laws for the logical connectives that one pleases. This is a principle of tolerance, only some kind of scientific utility should make you prefer one to the other, but one is completely free to choose. Of course, a choice of a different logic is a choice of a different language form.

Now, here we already have the notion of adopting a logic, which is what I directed my remarks against last time. As I said, I don't think you can adopt a logic. Quine also criticizes this point of view and for *the very same* reason I did. He said, as against Carnap and this kind of view, that one can't adopt a logic because if one tries and sets up the conventions for how one is going to operate, one needs already to use logic to deduce any consequences from the conventions, even to understand what these alleged conventions mean.

This is all very familiar as a criticism of Carnap. Somehow people haven't realized how deep this kind of issue cuts. It seems to me, as I said last time, obviously to go just as strongly against Quine's own statements that logical laws are just hypotheses within the system which we accept just like any other laws, because then, too, how is one going to deduce anything from them? I cannot for the life of me, see how he criticizes this earlier view and then presents an alternative which seems to me to be subject to *exactly the same difficulty*. (Princeton Seminar. My emphasis.)

But this seems somewhat unfair or too flattering to Quine, depending the light in which to look at it. Is the difficulty Quine sees in Carnap "exactly the same difficulty" Kripke sees in Quine? There are reasons to suspect Kripke is reading his own adoption problem in Quine's Carroll argument.

Several differences can be found between the adoption problem and Quine's argument. The latter, for example, makes a general point; namely, that every convention attempting to set up a

logical principle, whatever the principle may be, will be confronted with the same problem: the need to use logic to extract consequences from the convention. Whether the principle is a basic one or not is irrelevant, since the point will apply to every logical principle. Moreover, Quine doesn't say that the logical inference needed is precisely the *same* one that the convention is trying to set up. He just says that "logic" or "a logical inference" will be needed. Of course, given the generality of the conventions, it is hard to see how UI and MPP inferences could be avoided. But the point remains that Quine was making a general claim: all logical conventions presuppose logical inferences for their application to individual cases, and the logical principle established by the convention and the logical inference required need not coincide (in fact, in most cases they will *not* coincide).¹²⁸

¹²⁸ We should note, however, that in "Truth by Convention" Quine gives an 'alternative' formulation that seems to veer closer to the adoption problem:

In a word, the difficulty is that if logic is to proceed mediately from conventions, logic is needed for inferring logic from conventions. Alternatively, the difficulty which appears thus as a self-presupposition of doctrine can be framed as turning upon a self-presupposition of primitives. It is supposed that the if-idiom, the not-idiom, the every-idiom, and so on, mean nothing to us initially, and that we *adopt* the conventions (I)-(VII) by way of circumscribing their meaning; and the difficulty is that communication of (I)-(VII) themselves depends upon free use of those very idioms which we are attempting to circumscribe, and can succeed only if we are already conversant with the idioms. (Quine 1936: 97. My emphasis: Quine uses here the very word "adopt.")

Boghossian (1996: 217, fn. 36) says that it is not clear that the two arguments are equivalent. It seems to me that they are not. Apart from the obvious point that this formulation is given in terms of meanings while the others are not (which, at any rate, is not a minor difference), in this case we won't need just some logic to extract logic out of the conventions, but a grasp of the very same connective the convention is trying to circumscribe. And, as many have argued, such grasping could very well require mastery of some relevant rules or principles. For this reason, this formulation could be regarded as bringing Quine closer to the adoption problem. However, as I said before, the adoption problem is not clearly a problem about meanings. Quine's argument would also apply, moreover, to connectives like 'or,' which are apparently not subject to the adoption problem, or at least they are not obviously subject to it. It is interesting that Quine quietly drops this 'alternative' way of presenting the argument in the later formulations. I will follow his lead and do the same, leaving the discussion of meaning-related problems for chapter 6.

In contrast, the adoption problem is only supposed to affect some basic logical principles or rules, like UI or MPP. It arises when inferences that accord with a certain rule of inference or logical principle are needed to apply those very same principles or rules to individual cases. And it is for this reason that they cannot be ‘adopted.’

But the main difference between the adoption problem and Quine’s argument is that the latter is intended to show that Carnap’s attempt to offer an account of the necessity and *a prioricity* of all the logical principles by a conventional stipulation of their meaning fails, because inferential transitions not explained by means of conventions will be required to apply the conventions to particular cases. In contrast, when we look at Quine’s views on logic in the light of the adoption problem we observe, *not* that the logical principles cannot be empirically confirmed or revised in the way Quine proposes (though this is a corollary), but rather that the status that he assigns to the logical truths in the web and his conception of the nature of the inferential transitions is untenable.

Quine’s Carroll-argument seems to be closer to what we have called the ‘standard objection’ than to the adoption problem. In both cases the problem appears to be that the justification of the truths of logic, whether by empirical or conventional means, will presuppose inferential transitions that correspond to at least some of those truths. If this is correct, Quine could still be accused of failing to grasp that the problem he sees in Carnap affects his own view. So Kripke’s claim would be vindicated, but the shared problem won’t be the adoption problem.

That Quine simply did not see “how deep this kind of issue cuts,” as Kripke says, appears to follow from his attitude towards conventions. As we observed, he rejects the idea that the truth of the logical principles could be grounded on linguistic conventions, but he doesn’t seem to think there is any problem in claiming that the logical principles (and hence the corresponding

inferences) could originate in conventions. He doesn't seem to be aware that his own argument posed a deeper problem that affected not only the justification of the logical principles by conventional means but also their status as stipulations, statements, or hypotheses that we simply accept. And it is plausible to think that precisely because he didn't see this that he thought that the logical truths could "provide" the connections. In other words, he didn't realize that his own proposal for the empirical justification of logic was threatened, not just because it was open to circularity objections, but because in his own conception we wouldn't have a web to begin with.¹²⁹

Before moving on to Descartes, we have to ask the million-dollar question: will the die-hard Quinean accept Kripke's version of Carroll as a knock out argument? No true die-hard Quinean would give up without at least trying a rejoinder. He may want to concede that Quine did not distinguished the beliefs about logical rules or principles from the actual inferences as much as he should have, and perhaps he might even admit that we cannot simply assume that the principles or rules of inference are 'adopted' (that is, that it is merely in virtue of our having accepted them or formed beliefs about them that we are in position to perform inferential transitions in accordance with them). But he may also want to say that once this distinction and some clarifications are introduced, the Quinean model would be in position to avoid Kripke's objection. Consider what Devitt, our paradigm of a die-hard Quinean, has to say:

The links that hold the web of belief together reflect a set of rules that are part of "an evidential system" (Field, 1996, 1998). As a result of nature and nurture each person embodies such a system which governs the way she arrives at her beliefs about the world. A system must include dispositions to respond selectively to perceptual experiences and to infer according to certain rules. A likely example of

¹²⁹ Kripke doesn't distinguish between what we have called 'the justification problem' and the 'inference problem.' Perhaps for this reason he identifies Quine's argument with the adoption problem.

a rule is *modus ponens*. So, a person embodying an evidential system *S* containing this rule is disposed to infer according to the pattern:

If *p* then *q*,
p,
So, *q*.

[...] for a person using *S* to have justified beliefs, its rules have to be good ones. This is not to say that she must know the epistemological theory,

T: *S* is a good evidential system,

for her beliefs to be justified. So it is not to say that she must know

MP: *Modus ponens* is a valid inference

for her beliefs to be justified by *modus ponens* arguments. Indeed, as Lewis Carroll made clear a century ago, the demand for this sort of extra premise in an argument leads to a regress (see Boghossian, 2001, for a discussion). And it is just as well that our person is not required to have this epistemological knowledge for, if she is an ordinary member of the folk, she is unlikely to have given such matters much thought. Still, it is certainly appropriate *to* give them thought. So it is appropriate for the person to stand back from her arguments and ask some epistemological questions. What are the rules of *S*? This is a question in *descriptive* epistemology. Are the rules of *S* good? How do we know *T* and *MP*? These are questions in *normative* epistemology. Any answers to such questions will be further beliefs, additions to her web. (Devitt 2005: 108)

According to this, beliefs about the validity of a rule of inference or logical truth need not have a fundamental role in the Quinean model. Contrary to what Quine himself may have said, it is not the case that beliefs about logical truths or rules of inference “provide” the inferential connections of the system, but the other way around: we start out with an innate evidential system which is likely to include dispositions to infer according to rules such as MPP and UI; only later are we in position to identify inferential patterns and decide whether to include them in the web as new beliefs.

In Devitt’s view, we “embody” an innate evidential system, and our acquisition and evaluation of beliefs is governed by it. We form and evaluate beliefs even when we have no beliefs *about* the inferential patterns we are conforming to, even if we have not yet developed a *logica docens*.

Identifying and adding beliefs *about* the inferential connections as part of the web is something relatively new in human evolution (“after all,” Devitt wrote to me, “didn’t some ancient Greek (Aristotle?) *first* figure out a logical principle? Someone surely must have”), and only very few of us possess such beliefs.¹³⁰ Thus, “once we see this,” our die-hard Quinean will insist, “we can reject Kripke’s claim that the Quinean model presupposes that basic logical principles or rules of inference are ‘adopted.’” Devitt’s reply is then that Kripke’s charge may apply to Quine himself, but not to this version of the Quinean model in which the *logica utens* takes priority over the *logica docens*. We don’t adopt our *logica utens*, we come with it.

This rejoinder, however, seems insufficient as stated. As long as we continue to claim that in our practice of inferring according to MPP or UI patterns we somehow rely on the general content of such rules – perhaps because we possess a mental state that encodes the content of the rules and implicitly guides our applications – the problem will recur. Because, if that is the case, why isn’t Harry helped by our telling him the relevant rule or principle? Saying that we embody such rules and that the mental states encoding them are innate won’t make a difference. The situation may perhaps be different if we are willing to claim, as Devitt does, that the embodied rules are ‘hardwired’ in us and govern our inferences by means of a processing operation without being represented. Since hardwired rules will be discussed in chapter 7 (section 7.2.2), let’s assume that they manage to avoid the previous concern and constitute a plausible alternative. There remains, however, another worry (well, actually, a number of worries, since some of the

¹³⁰ This point is also stressed in his other papers on the subject. For example, Devitt (1998) says that “[...] the person who embodies *S* need not have any *theory* of these matters. So she need not *believe* that *modus ponens* is a valid inference nor that ‘If *p* then *q* and *p*, then *q*’ is a logical truth. Indeed, in my view, it is almost certain that she does not have such beliefs innately, and unlikely that she will ever come to have them, given our ordinary indifference to logical theory.” (1998: 266)

problems discussed in the previous section in connection to the standard objection will recur) concerning confirmation, revision, and the matter of adoption.

According to the view sketched above, we have to start out with whatever innate evidential system we happen to have – including whatever deductive rules are built in it – because we have nothing else to go by, because “we need logic to get evidence for or against anything.” (2011b: 277) But Devitt thinks that though we have no choice but to go by that system *before* any empirical evidence for or against it can be considered, the beliefs obtained by a processing operation governed by, say, MPP, will only be justified if the rule itself is valid. In his view:

... experience justifies beliefs in the interior of the web via logical links with beliefs at the periphery, via logical links with beliefs “close to experience”. But these justifications depend on the logical links themselves being justified: clearly a belief is not justified by other beliefs unless those others give it *genuine support*.” (2011b: 277)

The problem is, as he himself acknowledges,¹³¹ that we have no reason to suppose in advance that the innate deductive rules in *S* are ‘good.’ So, being governed by an innate evidential system that contains the MPP rule and that, at least initially, is not open for us to dispute, does not settle the issue of whether we should or not conform to the MPP pattern in figuring out what to believe.¹³² This depends on the empirical confirmation of the *belief* about the validity of MPP (together with the belief about the ‘goodness’ of *S* as a whole).

¹³¹ In his reply to Rey (1998), Devitt says: “The evidence suggests that some “good” and some “bad” logical sub-systems may be innate.”

¹³² He says:

A person must begin the pursuit of knowledge using some innate evidential system – a set of deductive and inductive rules – before she has any empirical evidence for its legitimacy. So, in that pragmatic respect, it is indeed “reasonable” for her to use this system [...] However, in another respect, most of us would hold that it remains to be seen whether it is reasonable for her

Our warrant for the *use* of inferential patterns that accord with MPP rests then with the empirical justification of the *belief* in its validity. Moreover, since “justified beliefs are produced and/or sustained by experiences in a way that is appropriately sensitive to the way the world is,” (2011b: 272)¹³³ it is to be expected that any belief, including any belief we may have about *S*, will also be empirically revisable. And Devitt does not disappoint: “all parts of *S* are empirically *revisable*.” (2011b: 279)

Yet, it is not at all clear in this model how revising a *belief* about, say, the validity of a basic rule, could lead to a revision of the inferential *practice* (the *logica utens*), especially if we take into account that the practice is generated by an innate embodied system in which the rules are not represented but built in (hardwired).

Devitt makes the following remarks on the matter of revision:

The Quinean empiricist view of logic is more explicitly concerned with *T* rather than *S*, with beliefs about logic rather than with logical practices. The claim is that these beliefs, although central in the web of belief, are nonetheless revisable in the light of experience. Still, this claim has clear consequences for *S*, the system that “holds the web together”. Suppose that experience leads a person to abandon *T* in favor of *T'*, a theory that recommends an evidential system *S'* built around a nonclassical logic. Then clearly the person should abandon *S* in favor of *S'*. In this way the Quinean thinks that our logical practices are themselves open to rational revision in the light of experience. (1998: 266)

The picture outlined here is open to objections related to the standard objection. We have already reviewed some of Devitt’s attempts to overcome it, so let’s ignore them for the moment.

to do so. In this other respect, we are concerned with the *justification* of *S*, with whether or not *S* is a *good* system, whatever the particular person may have reason to believe about this. (1998: 264)

¹³³ He claims that, “there is *only one* way of knowing, the empirical way that is the basis of science (whatever way that may be).” (1996: 2)

It should be noted that not only the empirical confirmation, but also the possibility of an empirical revision is central to Devitt's proposal. Otherwise, there would be little point in insisting that it is only because the beliefs about the deductive rules in S are well confirmed (that is, that T is well confirmed), that our inferential patterns *should* conform to them. If we had no means of disconfirming or revising such beliefs, we would have to employ whatever inferential patterns are the result of our innate hardwired processing operations in any case. For this reason, I don't think that Devitt is in position to claim that "perhaps it would be wiser to remain neutral on the matter of empirical revision, resting the naturalist case on the empirical justification of logic." (2011b: fn. 9) So, how is the revision supposed to occur?

Devitt doesn't give a detailed account, but it seems clear that changes in the inferential practice are the result of changes in the beliefs about the rules employed in the inferential practice. This, however, suggests that the initial priority of the *logica utens* was only temporary: it is in the end the *logica docens* what dictates what inferential patterns are permissible and what it is responsible for introducing modifications into the hardwired system.

What would happen if the belief in question concerns a basic rule like UI, MPP, or non-contradiction? It would seem that such revisions would require to suspend our usual way of inferring and 'adopt' a theory T^* that, say, allows some contradictions to be true, or that UI fails, which in turn is supposed to translate into a modification of the system S and hence the inferential practice it governs. But this is exactly the position Kripke is criticizing. What else could Devitt possibly mean when he says that we 'abandon' T in favor of T^* ?

Moreover, in Devitt's case it is particularly hard to see how this change in beliefs would translate into a change in S , since the rules in S are not represented but hardwired. On the representation model, it is relatively easy to see how a change would go, we change the

representation and that would (supposedly) change the practice (I'm not saying here that this would be possible, but that this is how the change would go if we indeed relied on representations), but on Devitt's model the rules are hardwired, and so the very structure that generates our evidential system would have to be changed. Devitt compares represented rules and hardwired rules with the software and the hardware of a computer, changing the software is relatively easy, but changing the hardware requires rebuilding the computer. So, how can a change in beliefs generate a change in the hardwired rules?

It seems to me that Devitt's proposal, with its sharp contrast between our beliefs about rules of inference and the hard-to-change hardwired rules themselves, does not fit the Quinean model very well – which, as Katz emphasizes, is essentially a model of belief revision. I suspect this reply won't go a long way in persuading the die-hard Quinean, but we *must* declare this round over now and move on. We will give it another shot in chapter 7. Meanwhile, we shouldn't get our hopes up; indeed, it looks as if the following might be analytic: die-hard Quineans are not very open to revision.

4.2 Descartes on Eternal Truths

Descartes famously thought that if the eternal truths were independent of God, they would constitute a limitation to his omnipotence.¹³⁴ God creates the eternal truths in the same way he

¹³⁴ Frankfurt (1977) points out that though the doctrine was presented in letters and also in the replies to the fifth and sixth objections to the *Meditations*, it was never included as an explicit part of his books. Descartes, however, originally intended to include it in a treatise on physics. In a letter to Mersenne he writes: "In my treatise on physics I shall discuss a number of metaphysical topics and especially the following. The mathematical truths which you call eternal have been laid down by God and depend on Him entirely no less than the rest of His creatures. Indeed to say that these truths are independent of God is to talk of Him as if He were Jupiter or Saturn and to subject Him to the Styx and the Fates." [April 15,

creates any other truth, and it is also him who sustains their existence. His creative power, furthermore, is not limited by anything; his will is free of external influences or reasons besides his own willing. Among eternal truths, in Descartes's view, are the following: "it is impossible that the same thing can be and not be at the same time," "what has been done cannot be undone," "he who thinks must exist while he thinks,"¹³⁵ and "the interior angles of a triangle are equal to two right angles."¹³⁶ He writes:

... it is impossible to imagine that anything is thought of in the divine intellect as good or true, or worthy of belief or action or omission, prior to the decision of the divine will to make it so. I am not speaking here of temporal priority: I mean that there is not even any priority of order, or nature, or of 'rationally determined reason' as they call it, such that God's idea of the good impelled him to choose one thing rather than another. For example, God did not will the creation of the world in time because he saw that it would be better this way than if he had created it from eternity, nor did he will that the three angles of a triangle should be equal to two right angles because he recognized that it could not be otherwise, and so on. On the contrary, it is because he willed to create the world in time that it is better this way than if he had created it from eternity; and it is because he willed that the three angles of a triangle should necessarily equal two right angles that this is true and cannot be otherwise; and so on in other cases. (*Sixth Set of Replies*: 291)

Against Thomas Aquinas and many other medieval thinkers who held that God could not have created a world which violated the laws of logic,¹³⁷ Descartes believed that just as God

1630, Kenny (1970), quoted by Frankfurt 1977: 36].

¹³⁵ The first three examples are listed in Descartes (1644, Principle 49: 238-239).

¹³⁶ Mentioned in the *Sixth Set of Replies* (Descartes 1641: 291), see below.

¹³⁷ Aquinas writes:

... all confess that God is omnipotent; but it seems difficult to explain in what His omnipotence precisely consists: for there may be doubt as to the precise meaning of the word 'all' when we say that God can do all things. [...] "God can do all things," is rightly understood to mean that God can do all things that are possible; and for this reason He is said to be omnipotent. Now according to the Philosopher (Metaph. v, 17), a thing is said to be possible in two ways. First in relation to some power, thus whatever is subject to human power is said to be possible to man. Secondly

could have decided to make true a different set of mathematical propositions, he could also have decided to make true a different set of logical principles. In a letter to Mesland he explicitly says that God could have made contradictions true: “God cannot have been determined to make it true that contradictories cannot be true together, and therefore . . . He could have done the opposite” [Letter to Mesland, May 2, 1644, Kenny (1970: 151).].¹³⁸

Now, in contrast to more contemporary worries about logical revisionism and the epistemic notion of the *a priori*, Descartes’s concern is metaphysical. His problem is whether God could have created a world in which the principles of logic do not hold, whether such a world could exist. And, according to Descartes, the answer to this question is, ‘yes; God could have decreed that such a world exist.’

absolutely, on account of the relation in which the very terms stand to each other.

[...] It remains therefore, that God is called omnipotent because He can do all things that are possible absolutely; which is the second way of saying a thing is possible.

[...] The divine existence, however, upon which the nature of power in God is founded, is infinite, and is not limited to any genus of being; but possesses within itself the perfection of all being. Whence, whatsoever has or can have the nature of being, is numbered among the absolutely possible things, in respect of which God is called omnipotent. Now nothing is opposed to the idea of being except non-being. Therefore, that which implies being and non-being at the same time is repugnant to the idea of an absolutely possible thing, within the scope of the divine omnipotence. For such cannot come under the divine omnipotence, not because of any defect in the power of God, but because it has not the nature of a feasible or possible thing. Therefore, everything that does not imply a contradiction in terms, is numbered amongst those possible things, in respect of which God is called omnipotent: whereas whatever implies contradiction does not come within the scope of divine omnipotence, because it cannot have the aspect of possibility. Hence it is better to say that such things cannot be done, than that God cannot do them. Nor is this contrary to the word of the angel, saying: “No word shall be impossible with God.” For whatever implies a contradiction cannot be a word, because no intellect can possibly conceive such a thing. (*Summa Theologiae*, Q. 25)

¹³⁸ Gary Ostertag pointed out to me that Descartes probably had a different conception of logic. He suggested that even if Descartes thought that there could be true instances of a contradiction (or that God could make them true if he willed so), it doesn’t follow that he was committed to an actual logical law that God could have changed. I’m not sure how this suggestion could be worked out: whatever Descartes’s conception of logic might have been, even if it did not include anything we would call a logical principle or law, such commitment *does* seem to follow. At any rate, this is not meant to be an exegetical piece on Descartes. Regardless of its historical accuracy it seems to me that the point remains of interest.

In opposition to Descartes's view, Kripke has this to say:¹³⁹

We do not need to consider whether the laws of logic are necessary by divine decree, or whether a coherent conception of their necessity could accord with this. It is enough to see whether a divine decree could make the laws of logic true (in the actual world). My claim is that God's decree that every universal statement implies each of its instances will be of no use.

For example, would the decree necessitate that "everything can be at most at one place at a given time" implies that such and such a rock can be in at most at one place at a given time? No, because according to Descartes there is a possible world in which every universal statement implies each of its instances, but at the same time a particular universal statement about places does not imply each of its instances. Similarly for any other particular universal statement. Thus, the decree by God that universal instantiation does hold in the world he created would be in and of itself useless.

To repeat: is it sufficient to explain why every universal statement implies its instances that God decreed that this be so? No, because it would be possible that God should have decreed that every universal statement implies its instances and also that a particular universal statement fails to imply all of its instances. Therefore, such a decree is ineffective by itself (and God of course could have made such a conflicting decree because, according to Descartes, he is not bound by the laws of logic or anything else).

The laws of logic need to hold for these divine decrees to be at all effective, otherwise God could decree that the three angles of a triangle equal two right angles, but nevertheless at the same time decree that the three angles of a particular triangle do not equal two right angles. Hence, such divine decrees are ineffective as to explain why the laws of logic hold in the actual world, let alone other possible worlds. With other cases, as for example why are there frogs, a decree could work as an explanation, but not in the case of logic.

God's decree cannot explain why certain basic logical principles hold in the actual world, and hence it cannot do the explanatory work it was meant to do.

Charges of incoherence against Descartes's doctrine of the creation of eternal truths have been raised before.¹⁴⁰ But Kripke raises a different complaint. He criticizes the decree not for being

¹³⁹ My thanks to Saul Kripke for producing a written statement of his argument to be included here.

¹⁴⁰ See, for example, Geach (1973: 10) and Nagel (1997: 60ff), who argues that Descartes's views are "unintelligible" (note that he claims to have been influenced by Kripke's 1974 Princeton Seminar). See also Frankfurt (1977) for a defense of Descartes.

incoherent or unintelligible, but for being useless. The objection is given in the form of a modal argument that, among other things, intends to bring out the uniqueness of the case of logic.

The argument could be thought as a variation of the adoption problem: in the case of basic logical principles, a divine decree would be, just as Harry's acceptance of basic logical principles or rules, without effect. Neither God nor Harry are bound by UI: the decree intends to make UI true but God's own will is not constrained by it or any other logical principle, and Harry does not infer in accordance with UI. Because of this, the attempts to either establish the truth of UI by means of a decree, or, in Harry's case, adopt UI, fail. Indeed, if God were trying to make sure UI held in his created universe on the basis of his own decree, he would fail for the same reasons Harry does. Alas, basic logic doesn't show much respect for God's omnipotence.

In contrast to Harry's case, however, non-contradiction plays a key role in Kripke's argument against Descartes. It is because God is not bound by non-contradiction that his decrees about other logical principles, like UI, turn out to be ineffective. Another point of divergence is that whereas the adoption problem makes an epistemological point, Kripke's modal argument is, following Descartes's concerns, metaphysical.

Perhaps an epistemological version of the argument could go along the following lines. Assume that in the actual world the usual basic principles of logic hold and that our minds cannot conceive a world in which they do not. Descartes might subscribe that such a world is inconceivable *for us*, but he would probably want to add that this is just a limitation of our minds. God's omnipotence demands that he be able to create a world in which those principles do not hold. Suppose now that Descartes's conception of a completely arbitrary divine will and unlimited power is consistent with the idea that God could reject, if he wished to do so, the

logical principles of the actual world.¹⁴¹ Imagine that God decrees the following: all inferences of the form ‘all universal statements imply all of their instances’ are fallacious. Let’s assume as well that before the new decree takes effect we inferred in accordance with UI and that God doesn’t change our inferential uses in any way. Our *logica utens* remains as it ever was. (We are presupposing here that on Descartes’ view of God’s unlimited power it makes sense to think that God could decree that the laws of logic change while leaving our actual inferential processes intact. But whether this is so is unclear.)¹⁴²

At any rate, were God to do both of these things, what would the effect of his decree be? The problem is that, for us, any application of the new decree requires what the decree itself precludes: inferring in accordance with UI. So the only way we could find God’s decree useful is by violating it, for otherwise it would be without effect.

Analogous arguments can be given for non-contradiction, MPP, and adjunction. Consider non-contradiction. The decree could be as follows: not all instances of “A and not A” are fallacious. But, to generate anything out of the decree, we will need to assume that non-contradiction applies to the decree itself, otherwise we would find it useless. In the case of MPP,

¹⁴¹ It is not clear whether, for Descartes, God can change the logical principles once they have been created or would be bound by them. In a letter to Mersenne he writes:

It will be said that if God had established these truths He could change them as a king changes his laws. To this the answer is ‘Yes, he can, if His will can change.’ ‘But I understand them to be eternal and unchangeable’ – ‘I make the same judgment about God’ ‘But His will is free’ – ‘Yes, but His power is incomprehensible.’ [Letter to Mersenne, April 15, 1630 (Kenny 1970)]

The passage is, to my mind, not very clear, and Descartes appears to be hesitant. On the one hand, one could argue that if God is to have real unlimited power to create the principles of logic and sustain them, he should be able to change them too. On the other hand, it may be argued that the situation imagined in the argument above can be ruled out because it would either make God a deceiver or betray a flaw. Descartes’s doctrine of the eternal truths has puzzled commentators for centuries and I won’t try to clarify it any further here. If the epistemological argument is sound, it will be of some interest whether it can be applied to Descartes’s views or not.

¹⁴² Here, again, it is likely that Descartes thought that a change in the laws of logic, if possible, would bring about a change in our minds (see fn. 138).

the decree would be: all inferences of the form ‘if A implies B, and A holds, then B follows’ are fallacious. Here too, to pronounce an instance of MPP fallacious by reference to the new decree, we will need to perform a MPP inference: if any inference is an inference of the form ‘if A implies B, and A holds, then B follows,’ then that inference is fallacious. Here we have an inference of the form ‘If A implies B, and A holds, then B follows.’ Therefore, this inference is fallacious. For the case of adjunction, the decree will be that “all inferences of the form ‘A, B, then A and B’ are fallacious.” But we need to put together the decree along with an instance of an adjunction inference to conclude that it is fallacious, and adjunction is needed to put the two premises together. Notice that, if instead of following UI, we were rather following an alternative principle like PInst (perverse instantiation), God’s corresponding decree would not help either. We would be in Ernie’s situation.

So God’s decrees rejecting our basic logical principles create a situation in which, if we try to follow them, we have to violate them, and if we refuse to violate the decrees, we cannot apply them and they are without effect. Moreover, if God were not trying to revise basic logical principles by decree but to make us infer in accordance with them (without generating past uses or altering our minds), his decrees wouldn’t be of use if we didn’t already infer in accordance with them. We would be in a situation similar to Harry’s. Thus, leaving aside Kripke’s metaphysical argument and assuming that God could create basic logical principles, for the decrees to be of any effect more than that is needed: we need to already infer in accordance with the principles in question.¹⁴³

¹⁴³ Michael Levin has suggested to me a further way to put the ‘without effect’ argument. If God decreed that the laws of logic did not hold of the universe (or no longer held of the universe), what would the universe be like? If we use the laws of logic to figure out what the universe would be like, we are violating the terms of the decree. If we do not use the laws of logic to figure out what the universe would be like, then, of course, we cannot figure out what the universe would be like.

In my epistemological rendering of the argument, the appeal to non-contradiction is not needed. The argument can be applied to non-contradiction (either in the rejection or getting-us-to-reason mode) as it can be applied to other basic principles, but it doesn't presuppose that the decrees are ineffective because non-contradiction fails.

One final point is worth mentioning. Arnold Koslow has suggested that, in view of the case of perverse instantiation and the argument against Descartes, maybe the adoption problem can also be formulated in terms of rejection. Let's call it the rejection problem, and let's formulate it, in the light of our prior discussion, as follows:

Rejection Problem (RP): certain basic logical principles cannot be *rejected* because, if a subject already reasons in accordance with them, rejecting them will require their use, and if the subject does *not* reason in accordance with them, no rejection is possible.

Let me make clear that we are understanding "rejection" in terms of the acceptance of a statement or decree that holds that "all inferences of the *X* form are fallacious" (where the inferences in question are basic, like UI or MPP). There are, of course, other ways in which a logical principle could be rejected. It would be extremely implausible for anybody to claim that all inferences of the MPP or UI form are fallacious. Indeed, it is doubtful that we could even make sense of such a decree.

The rejection problem *should not* then be read as claiming that basic logical principles cannot be revised.¹⁴⁴ Its claim is much more modest: that a particular way of “rejecting” them – one which is based on the acceptance of statements or decrees which affirm that “all inferences of the *X* form are fallacious” – won’t work for some basic logical principles.

¹⁴⁴ Similarly, Kripke says that he does not take his criticisms of Quine to show that logical principles are unrevisable. If Quine had not claimed that logical principles are open to revision, the “Two Dogmas” conception of logical principles as mere statements or beliefs would have been enough to make it objectionable from the perspective of the adoption problem. But this is a separate issue. There may well be other ways of defending revisionism that do not presuppose such a conception and that are not thus subject to the adoption problem.

Chapter 5

Kripke on Logic and Intuition

True to his style, Kripke leaves us in perplexity without offering a clear account of his own to answer the challenge he has identified. To be fair, though, we should remember that the main target of Kripke's seminars and lectures was the view that we could *adopt* alternative logics. As a way of approaching the issue he discusses whether we may be able to adopt basic logical principles or rules, like UI or MPP. But once he rejects this possibility his main focus is logical revisionism via Putnam's proposal, and not so much whether and how we could justify such principles or rules, or how to understand the corresponding inferential transitions. He suggests, nevertheless, a couple of different, maybe even conflicting, lines of thought.

The (apparent) conflict was mentioned in chapter 3 (section 3.3.3). We said there that though it is already present in *WRPL* (see fn. 90), the tension becomes evident in the lectures on the nature of logic, since in this case Kripke is speaking in his own name. At times he seems to be attracted to what in *WRPL* he would call a 'Platonist' view, which takes intuitions to play a fundamental role by emphasizing either the need to *see* that a basic logical principle is self-evident, or to *see* that a particular conclusion follows. But we also find, pulling in what seems to be an opposite direction, a more 'Wittgensteinian' outlook in these lectures, where the importance of being in the *habit* of inferring in accordance with a pattern is stressed: at bottom what matters is the regularity of the inferential practice. Kripke makes passing remarks about these two 'outlooks' at different points, though neither is approached in a systematic way and their at least apparent incompatibility goes unnoticed. All this of course makes the reconstruction

of his views rather difficult and tentative. In this chapter we will focus on the Platonist outlook, leaving the discussion of the Wittgensteinian view for chapter 7. But before starting it would be good to know in advance that conversations with Kripke indicate that, to Wittgenstein's chagrin, the 'Platonist outlook' has long won the battle. In fact, he now seems to find his own remarks on inferential habits a little disturbing.¹⁴⁵ (They should, no doubt, be attributed to his alter-ego, who appears to have been quite active in the 70s.)

5.1 Self-Evidence and the Analogy between Formal Systems and Non-Euclidean Geometries

What is, then, the role Kripke assigns to intuitions? The answer is not entirely clear. Every now and then he just seems to be talking about our usual ability or capacity to infer, our reasoning intuitively in accord with, say, the UI rule. He claims, for example, that intuition cannot be "thrown out to the dogs," because we either "reason intuitively or not at all." But though he talks of 'intuitions' and 'intuitive reasoning,' the context often suggests that he could call it a 'natural' or 'usual' way of reasoning without any serious modification of what he intends to say. In fact, if this is all he means, it may be a good idea to avoid the term 'intuitions'

¹⁴⁵ Not surprisingly, the 'Wittgensteinian outlook' is much more explicit in the Princeton Seminar (the seminar meant to be devoted to Wittgenstein). Though Kripke introduces his remarks on habits in his own name, it is clear from some expressions he uses in these lectures that he was reading Wittgenstein's *Remarks* at the time (especially Part I).

We should keep in mind that Kripke has not edited these lectures and seminars. It is not hard to imagine some slips that would have been edited out had the lectures been corrected. I have tried to focus on the points he repeats at different times. I don't think, however, that the 'Wittgensteinian outlook' was just a slip. On the contrary, it seems clear to me that it was a view that Kripke found compelling (at least then).

altogether, given the suspicion it tends to generate.¹⁴⁶ Every so often, however, he also speaks about ‘seeing’ that a principle or inference is self-evident, suggesting that intuitions are doing substantial work after all.

I won’t be discussing in this chapter Kripke’s use of ‘intuitions’ as just an intuitive (or, as suggested, ‘natural’ or ‘usual’) ability to infer; that way of speaking is related to the role he assigns to habits, to be discussed in chapter 7. I will focus instead on his more substantive use of the notion of intuition. Unfortunately, this doesn’t necessarily simplify the reconstruction of his views, since here, too, Kripke is somewhat ambiguous: while some passages suggest that *seeing* the logical principles (or the validity of the rules) with some kind of intuitive self-evidence is the key issue, others seem to stress that we need to *see* that conclusion follows and hint at an intuitive grasp of a particular conclusion. To try to make sense of this we will have to go over Kripke’s arguments and take into account the background of his remarks on intuitions. A bit of patience will thus be required.

¹⁴⁶ The role of intuitions in philosophical theorizing has been subject to scrutiny for some time now. Though some reject the legitimacy of appeals to intuitions in philosophical discussions (see, for example, Kornblith 1998, Weinberg 2007, Weinberg *et. al.* 2010, and Carruthers 2009a and 2009b), they seem to be widespread. However, as Williamson points out, “there is no agreed or even popular account of how intuition works, no accepted explanation of the hoped-for correlation between our having an intuition that P and its being true that P.” (Williamson 2007: 215). One of the problems seems to be that even when talk of “intuitions” is regarded as in need of explanation, views held on issues such as *a priori* knowledge, introspection, consciousness, naturalism, etc. are likely to shape the accounts offered and the role assigned to intuitions in philosophical arguments. It is, for example, hardly surprising to find Devitt defending a “naturalistic and non-Cartesian view of intuitions in general.” According to him, “they are empirical central-processor responses to phenomena differing from other such responses only in being immediate and fairly unreflective.” (2006b: 293) Their role, far from delivering *a priori* knowledge, is to “reflect an empirically based expertise in the identification of kinds” (2011b: 278) In contrast, Bealer defends a Cartesian view of intuitions as “irreducible” and “natural” propositional attitudes. For him, “intuition is an “intellectual seeming”” (1998: 207. See also 2002: 73) and the role of philosophy is to provide a systematization of those intuitions that have survived a “dialectical critique” until an equilibrium has been reached. (See Rey 2013 for a criticism of Devitt’s and Bealer’s views.)

It is worth pointing out that Cappelen has recently challenged the claim – which he calls the “centrality thesis” – that analytic philosophers do in fact rely on intuitions as evidence for their views. See Cappelen 2012.

Now, Kripke tends to introduce references to intuitions in the context of his complaints about formal systems. We observed (pp. 97-98) that he thinks that the existence of alternative formal systems has wrongly led us to believe that the acceptance of logical principles or rules is all that is needed to produce derivations. In a similar passage, this time from the Pittsburgh lecture, he says:

The whole illusion that has created all of this talk about adopting one logic and so on, as if one could do it, basically comes from this. We are all taught in logic textbooks that when you have a formal system you do not need to reason. A machine can do it, a moron can follow these directions; all you have to be is programmed, and so on. This of course is true; a machine can do it, all it has to be is programmed. Then one thinks that one has something really solid that is independent of any reasoning; *just do this*. (Pittsburgh Lecture)

According to Kripke, one of the things that give support to this way of thinking – which he calls the ‘formal system fallacy’ – is the alleged analogy between logic and non-Euclidean geometries:

Just as in the case of non-Euclidean geometry we throw intuition to the wind and adopt an axiomatic system as supposedly describing the real physical world, in which some situation exists where parallel lines can somehow sneak in and if not meet, approach each other, get closer and closer, so in every other domain we cannot rely on intuition. Once one has a rival system of axioms, the mere fact that an old system of axioms struck us as the only intuitively acceptable should be given little weight. Once alternative geometries are under consideration, we abandon any preference – any near intuitive preference – for Euclidean geometry. Once alternative logics are under consideration we abandon any near intuitive preference for a particular system of logic. And so if one had a theory of marriage which implied that some – no doubt, only a few – of the married men are still bachelors, which satisfied some valuable criteria of internal coherence and predictive power – and it probably would have considerable predictive power – then we might consider it as an alternative to the present system and throw intuition here to the wind. (Pittsburgh Lecture)

The target of this passage is, once again, Putnam (1968).¹⁴⁷ In this paper Putnam argues that a proper account of quantum mechanics – a ‘realist’ one, in his view – calls for the adoption of quantum logic, and hence provides a case for the empirical revision of logic. The details of the proposal need not concern us; it will be sufficient to know that the particular version of quantum logic Putnam favors, due to Birkhoff and von Neumann, requires the abandonment of the distributive law.¹⁴⁸

In his effort to make his case compelling, Putnam explicitly appeals to the case of geometry. He thinks there is a lesson of “philosophical importance” to be learned from its history; in particular, what interests him is the troubled history surrounding the fifth postulate of Euclidean geometry and the implications that, in his view, can be drawn for the notions of self-evidence and

¹⁴⁷ Kripke is alluding to the following passage:

If one did not know anything about non-Euclidean geometry, relativity, etc., would the intuitive evidence that this [that two lines *AB* and *CD* can have a constant distant apart over some interval and then begin to converge without bending] is a contradiction, an impossibility, a complete absurdity, etc., be any less than the intuitive evidence that no surface can be scarlet (all over) and bright green at the same time? Or that no bachelor is married? Is the intuitive ‘feeling’ of contradiction really different in the following three cases?

- (1) Someone claims that *AB* and *CD* are both *straight lines*, notwithstanding their anomalous behavior.
- (2) Someone claims that he has a sheet of paper which is scarlet (all over, on both sides) and green (all over, on both sides) at once.
- (3) Someone claims that some men are married (legally, at the present time) and nonetheless still *bachelors*.

It seems to me that it is not. (Putnam 1968: 174-175)

¹⁴⁸ See Birkhoff and von Neumann (1936). The ‘adoption’ of quantum logic is motivated by the anomalies associated with quantum measurements. Putnam defends a realist explanation of the famous two-slit experiment, which according to him involves giving up the distributive law (1968:180ff). He writes: “The only laws of classical logic that are given up in quantum logic are distributive laws, e.g. $p.(q \vee r) \equiv p.q \vee p.r$; and every single anomaly vanishes once we give these up.” (p. 184) (For a defense of quantum logic see also Finkelstein (1964)). There are, of course, other famous attempts to make sense of the anomalies of quantum mechanics that do not involve a “change in logic,” like the Copenhagen interpretation (criticized by Putnam), Bohmian mechanics, theories of the spontaneous collapse of the wave function, the many-worlds interpretation of quantum mechanics, etc.

necessary truth. As everyone knows – and the deniers of the *a priori* won't ever let anyone forget – the Euclidean system was for a long time considered a paradigmatic example of *a priori* knowledge, singled out by Kant as an example of the synthetic *a priori*. Nevertheless, almost from the beginning, and maybe even to Euclid's own mind, the fifth postulate about parallel lines was found to be more complicated and less evident than the other four.¹⁴⁹ This is hardly surprising considering Euclid's original formulation: “that, if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles.” (*Elements*, Book I) There were many attempts to make it simpler and more evident, but, even though some alternative formulations made it look less convoluted, doubts persisted nonetheless and the fifth postulate continued to worry mathematicians.

In the 1820s, probably motivated by the dissatisfaction that the ‘fifth’ had long been generating, Lobachevski and Bolyai – and maybe even before, Gauss, though he did not publish his results – independently discovered what is now known as Bolyai-Lobachevsky non-Euclidean geometry (or hyperbolic geometry). The problematic fifth postulate was replaced with the assumption that through a point not on the line there exists more than one parallel to the line. Later on, in the 1850s, Riemann came up with another way of negating the ‘fifth’ that assumed that any two straight lines in a plane intersect. He also modified the remaining postulates to

¹⁴⁹ The Euclidean system comprises five postulates, twenty-three definitions, and some common notions. That the fifth postulate was considered less straightforward than the other four gives, in Kripke's opinion, some pause to the claim that a self-evident principle was found to be a falsehood. He says:

Why people started to investigate non-Euclidean geometry to begin with? It was because the parallel postulate was thought not to have been as self-evident as the others. This is starkly true. And they tried to prove it from the others and then they found, finally by exhibiting non-Euclidean [...], that it couldn't be proved. So that means that there must have been some kind of doubt as to its self-evidence *all along*. (Princeton Seminar)

preserve consistency and obtained what is now called Riemannian or elliptical geometry.¹⁵⁰ One very simple way of understanding Riemannian geometry is to think of a sphere where for any line L and point A not on L , all lines going through A intersect L ; a “line” or “geodesic” in this model is specified as the arc of a great circle.

The discovery wasn’t considered all that significant outside mathematics in the beginning; non-Euclidean geometries were for decades mostly regarded as capturing geometrical fantasies with no physical application. Some, most notably Frege, who sided with Kant in claiming that Euclidean geometry rests on an *a priori* intuition and gives us *a priori* knowledge of physical space, simply dismissed them.¹⁵¹ But most changed their attitude when Einstein’s theory of general relativity made its appearance and a physical application for Riemann’s geometry was found. General relativity establishes a connection between mass and space in which the curvature of space varies according to the concentration of matter in the region. As a consequence of the

¹⁵⁰ Riemannian geometry is in fact an infinite family of non-Euclidean geometries.

¹⁵¹ In his most severe moods Frege compares non-Euclidean geometries to alchemy and astrology and seems to think that “non-Euclidean geometry will have to be counted amongst the pseudo-sciences, to the study of which we still attach some slight importance, but only as historical curiosities.” (Frege 1899-1906: 168. My thanks to Kripke for pointing out this passage to me.) However, in *The Foundations of Arithmetic* he writes in a much more conciliatory tone: “... the truths of geometry govern all that is spatially intuitable, whether actual or product of our fancy. The wildest versions of delirium, the boldest inventions of legend and poetry, where animals speak and stars stand still, where men are turned to stone and trees turn into men, where the drowning haul themselves up out of swamps by their own topknots – all these remain, so long as they remain intuitable, still subject to the axioms of geometry. Conceptual thought alone can after a fashion shake off this yoke, when it assumes, say, a space of four dimensions or positive curvature. To study such conceptions is not useless by any means; but it is to leave the ground of intuition entirely behind. If we do make use of intuition even here, as an aid, it is still the same old intuition of Euclidean space, the only one whose structures we can intuit. Only then the intuition is not taken at its face value, but as symbolic of something else: for example, we call straight or plane what we actually intuit as curved. For purposes of conceptual thought we can always assume the contrary of some one or other of the geometrical axioms, without involving ourselves in any self-contradictions when we proceed to our deductions, despite the conflict between our assumptions and our intuition. The fact that this is possible shows that the axioms of geometry are independent of one another and of the primitive laws of logic, and consequently are synthetic.” (Frege 1884: 20)

positive curvature predicted by the theory, the resulting physical space cannot be described by Euclidean geometry and a version of Riemannian geometry is used instead.¹⁵²

For Putnam, Euclidean geometry is “a case in which ‘necessary’ truths (or rather ‘truths’) turned out to be falsehoods.” (p. 174) It is, in his view, a clear intuitive contradiction that two straight lines for which the distance apart is constant over some interval, can start to converge after some point without bending at all.¹⁵³ Just like something being both red and green all over, or married men being bachelors (see fn. 147 above). Nevertheless, “what was yesterday’s ‘evident’ impossibility is today’s possibility (and even *actuality* [...])” (p. 175). And of course, the main point he ultimately wants to make is that exactly the same could happen in the case of logic.

He views “the analogy between the epistemological situation in logic and the epistemological situation in geometry as a perfect one” (p. 190), and claims that if his interpretation of quantum mechanics is right, then “... *logic is as empirical as geometry*. It makes as much sense to speak of ‘physical logic’ as of ‘physical geometry’. We live in a world with a non-classical logic. Certain statements – just the one we encounter in daily life – *do* obey classical logic,” but according to him “... Quantum mechanics itself explains the *approximate* validity of *classical* logic ‘in the large’, just as non-Euclidean geometry explains the *approximate* validity of *Euclidean* geometry ‘in the small’.” (p. 184)

¹⁵² For the historical details, I mainly follow Sklar (1974).

¹⁵³ Putnam rejects the idea that ‘straight line’ and ‘geodesic’ have a different meaning. One of the reasons he gives is that if we make such a claim, we would also have to reject that ‘a straight line is the shortest path,’ since this is preserved by the notion of a geodesic. He says:

The important point is that [‘straight line’] does not ‘change meaning’ in the trivial way one might at first suspect. Once one appreciates that something that was formerly literally unimaginable has indeed happened, then one also appreciates that the usual ‘linguistic’ moves only help to distort the nature of the discovery and not to clarify it. (1968: 177)

The idea seems to be that in the same way that Euclidean geometry was shown not to be *a priori* once non-Euclidean geometry was found to have a physical application, classical logic can be shown not to be *a priori* once alternative logics are found to have a physical application:

From the classical point of view, all this is nonsense, of course, since no empirical proposition could literally be more likely than that classical logic is right. But from the classical point of view, no empirical proposition could be more likely than that straight lines could not behave as depicted in Figure I [a figure showing two lines, *AB* and *CD*, which maintain a constant distance apart over some interval and then begin to approach each other]. What the classical point of view overlooks is that the *a prioricity* of logic and geometry vanishes as soon as *alternative logics* and *alternative geometries* begin to have serious physical applications. (1975a: 189) ?

Two points should be stressed here. First, Putnam does not distinguish in this paper between the metaphysical notion of necessary truth and the epistemological notions of *a prioricity* and self-evidence. Nevertheless, Kripke says that there is no big difference between Putnam and him in this respect, we just have to “read the terminus ‘necessary’ epistemically in every case in his paper” (Princeton Seminar). Second, what Putnam appears to be saying is not that a proposition that *seemed* to us to be self-evident turned out false, but that a proposition that continues to be as self-evident to us as it was before, was shown to be a falsehood. The lesson he wishes to extract from the case of geometry is therefore general: we cannot trust our own intuitions of self-evidence.

The whole category of ‘necessary truth’ is called into question. Perhaps it is for this reason that philosophers have been driven to such peculiar accounts of the foundations of the GTR [General Theory of Relativity]: that they could not believe that the obvious account, the account according to which ‘conceptual revolutions’ can overthrow even ‘necessary truth’ could possibly be correct. But it *is* correct, and it is high time we began to live with this fact.

And in the Preface to Putnam (1975b), he writes:

I was driven to the conclusion that there was such a thing as the overthrow of a proposition that was once *a priori* (or that once had the status of what we call “*a priori*” truth). If it could be rational to give up claims as self-evident as the geometrical proposition just mentioned, then, it seemed to me that there was no basis for maintaining that there are *any absolutely a priori truths*, any truths that a rational man is *forbidden* to even doubt. (1975b: xvi)

All the main claims Putnam advances are strongly rejected by Kripke: (1) that the analogy between logic and geometry is a perfect one; (2) that the case of non-Euclidean geometries can be used to draw support for logical revisionism; and (3) that the case of geometry shows that there are no self-evident truths. One of the reasons for (2) being a misguided strategy is that it overlooks the special nature of logic, which in turn makes the analogy between logic and geometry far from perfect:

Logic, even if one tried to throw intuition to the wind, cannot be like geometry, because one cannot adopt the logical laws as hypothesis and draw their consequences. You need logic in order to draw these consequences. There can be no neutral ground in which to discuss the drawing of consequences independently of logic itself. If there were it would be some special kind, called ‘*ur*-logic,’ and that would be what was really logic. (Pittsburgh Lecture)

Similarly, in the Princeton Seminar, he says:

Putnam claims that the analogy between logic and geometry is perfect. Now there may be a considerable analogy, I don’t know. But there is at least the following point of dis-analogy. Many people believed that the laws of Euclidean geometry were *a priori*, in contrast to Putnam, and probably in contrast to most people today. But whether they believed that or not, they thought they could deduce interesting consequences from the laws of geometry. Now in the case of some at least of the laws of logic, one could not deduce any consequences from them

unless one was in the habit of using the very pattern of inference that one is supposed to be assuming *already*. (Princeton Seminar) [Note the use of the word ‘habit’ here.]

There are, then, good reasons to think that the analogy between the case of logic and that of geometry is not as solid as Putnam is leading us to believe. When it comes to logic we cannot bypass intuitive reasoning by setting up a formal system because the very possibility of understanding and using the system depends on it. The existence of different formal systems and the fact that machines can deliver correct results in no way shows that intuitive reasoning is irrelevant. Kripke insists that we cannot either check the results that machines produce, or deduce anything from a formal system, without using “intuitive reasoning.”¹⁵⁴ In contrast, in the case of the different geometries we can just accept them as an abstract system of postulates and draw consequences (recall the case the case of Socrates’s slave boy discussed in fn. 61). Entities such as lines and points can simply be treated as abstract entities that behave in accordance with a certain set of postulates but lack any intuitive appeal. Kripke says:

You cannot undermine intuitive reasoning in the case of logic and try to get everything on a much more rigorous basis. One has to just think not in terms of some formal set of postulates, but intuitively – one has to reason. One cannot just adopt a formal system independently of any reasoning about it, because, if one tried to do so, one would not understand the directions for setting up the system itself. And so any comparison of logic to geometry, which says that in the case of

¹⁵⁴ Field gives a similar argument to question the analogy. He argues that “there is a fundamental difference in kind between logic and geometry” (1996: 361), because “... one needs logic in reasoning about anything (including geometry). By contrast, there are lots of things (logic is one of them) that one doesn’t need geometry in reasoning about; this asymmetry is certainly part of the explanation of why there is no obvious alternative to strong a priorism in logic.” (Field 1996: 369. See also Field 1998: 17ff) Even Priest, who defends the analogy between logic and geometry, concedes to Rescher (1969: 219) that “the articulation of a logical system requires the employment of a metalogic (normally an informal one, and not necessarily classical), whereas the articulation of a geometric system does not require the employment of a metageometry,” (Priest 2003: 443). But he doesn’t think that this asymmetry has a “significant import.”

logic, as in the supposed case of geometry, intuition can be thrown to the dogs, that is, [the suggestion that] reasoning outside the system of postulates can be thrown to the dogs, must be wrong. One can only reason as we always did, independently of any special set of rules called 'logic,' in setting up a formal system or in doing anything else. And if proof by cases was part of our intuitive apparatus, there is no analogy with geometry which says that this is not to be respected. (Pittsburgh Lecture)

For Kripke, the main moral that should be drawn from all this is the following:

There is no bedrock in the notion of a formal system which we can choose somehow to adopt or reject at will and then tailor one's logic to it. One has to first use reasoning in order to even see what is provable in a formal system. Formal systems don't say anything as they are; they are just formal objects. One reasons about them using the same reasoning that one always did. If one then gives them an interpretation, one can check, using the reasoning that we always did (and what other reasoning do we have? *None*, as I'm trying to say) whether under this interpretation they are (a) sound and (b) complete. (Pittsburgh Lecture)

Now, in the passages we have quoted so far Kripke doesn't appear to assign any crucial role to intuitions; he complains about the idea of "throwing intuition to the wind" and talks about "intuitive reasoning," but as we said before nothing of importance would seem to be lost by simply emphasizing that we need to exercise our 'natural' or 'usual' capacity to reason. Furthermore, these passages could be read as part of what we called the 'Wittgensteinian outlook,' since the main emphasis is placed on the inferential practice, on being in the 'habit' of reasoning in accordance with certain patterns of inference. Kripke himself seems to support this reading:

Pay no heed, so I say, to someone who says: "Forget your intuitions. Don't reason intuitively. Intuition here is question-begging term and also has a sort of undermining air. Well look, intuition, that's something that irrational people rely on. Saying you rely on intuition is not relying on anything very scientific" or

something like that. What really you've been asked to do is forget reasoning. Reasoning *is* the very process of just saying, "Look, this conclusion follows from this. If you accept that everything is green, it just follows that this is green." And if that's unscientific, I don't know what is scientific. But anyway if it were unscientific, there is nothing that we can replace it by; nothing more basic to which one can appeal. If you try to set up something more basic, say, a set of syntactic conventions, or a set of semantic conventions, or anything, you will find yourself, if you look really closely, reasoning in the same ordinary way as you set up these alleged conventions. (Princeton Seminar)¹⁵⁵

So "intuitive reasoning" is simply "reasoning," the pre-formal reasoning that we had all along. But this is not the whole of the matter, because Kripke also connects this "intuitive reasoning" with the notion of self-evidence, and it is here and in his defense of this notion from Putnam's attack that the 'Platonist outlook' emerges. Such connection can be appreciated in Kripke's discussion of the comparison between geometry and logic in terms of a blind and completely paralyzed man and someone that is "empty headed":

Euclidean geometry was thought to be *a priori*. Now, to a blind man at least the parallel postulate, that two lines perpendicular to the same line never meet, might not be self-evident. To see this as self-evident one has to do a little visualization; that's why Kant attached it to some form of sensuous intuition. [...] Of course he can try and trace it along as his hands, but supposing he had not tactile sense, or was very paralyzed, or something like that, it probably wouldn't be self-evident to him. He could still understand the axioms of geometry as an abstract system of postulates about some entity or other. The term "line" doesn't mean anything to him; it's just some kind of entity or a set of points; the term "point" doesn't mean anything in particular. Now that's the way geometry is conceived of in mathematics today. It's just a set of axioms about some kinds of entities, not necessarily lines and points in the intuitive sense. Then it is easy enough to see

¹⁵⁵ Here is a parallel passage from the Pittsburgh Seminar:

I am saying: one can't give up intuition, one can only reason intuitively. But that is to say, 'intuitively' is usually now a pejorative term, as in these papers. You know: "What's intuition? That sounds more like faith." On the contrary, I mean something like reason, one can only reason *by reasoning*. One cannot reason by setting up a formal system and then deciding to reason within it. Because, if one had no way of reasoning before, and tried to hold that in suspense, one could not understand this alleged setting up of a formal system. (Pittsburgh Lecture)

what an alternative geometry would be like. We can think *this way*. Then what we are told to do by some physicists is actually to apply this to the physical world, to ordinary lines and points. This is where the conflict with intuitions may come in. But to think about this, one can just imagine that one were a blind man, or didn't know what these entities were – lines and points and so on. One, so to speak, suspends one's visuality here. [...] Now the analogy to regarding oneself as if one were blind, or as if one didn't understand the meanings of the terms “line” and “point” and “angle” and so on, would be to regard oneself as if one were ... well, what's the analogue for logic of being blind? Empty-headed, I suppose. Or as if one didn't understand the logical particles, “and,” “or,” and “not.” But if you did that, you would be unable to think at all, and so you would be, in particular, unable to understand what consequences follow from this supposed alternative logic that one is going to adopt. (Princeton Seminar)¹⁵⁶

One way of reading the contrast Kripke is trying to set up here is to say that, while in the case of geometry we can reason within a system without finding its postulates self-evident, in the case of logic our “intuitive reasoning” (or just “reasoning”) according to some basic rules or principles hinges on finding the corresponding rules or principles self-evident. On this reading, it is because we find the UI rule self-evident that we find particular UI inferences evident as well and we reason “intuitively” according to them.

¹⁵⁶ Later on, he says:

... the analogy with geometry doesn't hold anyway, and this is what I have been trying to emphasize. In the case geometry, at least one can try this device of dis-interpretation. Even if one didn't understand how to interpret these things as points and lines, you can consider these hypotheses and draw consequences from them. This is what I'm arguing is not the case with respect to logic. So even if Putnam established that self-evident truths can be given up here, that would have little to do with the logical case, because the logical case is not a case of our accepting truths. If it were a case of accepting truths, then we could talk about giving them up. Maybe the case of Euclidean geometry *was* one of accepting truths, but the case of logic is not one of accepting truths. If one just accepted these truths, they wouldn't do you any good. A blind man who had no knowledge of geometry could be taught it as an axiomatic system, I suppose, and he could just accept these truths and draw consequences from them. He could even be told that people who can see believe these statements on some interpretation, but he could just deal with them abstractly. We can't do this in the case of logic – or at least it seems that we cannot. (Princeton Seminar)

Another possible reading is that it is not the principles or rules what we see as self-evident, but the particular inferences themselves. According to this, being empty headed or unable to understand the meaning of, say, ‘all’ is simply being unable to *see* that “this is green” follows from “Everything is green,” as Kripke suggests in one of the passages quoted before.

The first reading makes the parallel with the case of geometry straightforward, but the second reading seems to fit other passages and the general theme of these lectures better by giving the priority to the reasoning practice (the *logica utens*). Note that on both readings the subject is unable to “deduce anything from a formal system” because he lacks the relevant intuition of self-evidence. In Harry’s terms, we can say that his inability to reason according to UI is either explained by his not finding the UI principle or rule self-evident, or by his not finding particular inferences, such as that “Alex the lion talks” follows from “all the animals in the movie Madagascar talk” self-evident. In either case it seems clear that Kripke thinks that we cannot do without our intuitions of self-evidence for the reason that such intuitions are needed for “intuitive reasoning” to be possible.

Kripke’s defense of the notion of self-evidence from Putnam’s attack may help clarify his views. He complains that Putnam’s appeal to the case of geometry is meant as a “softening up procedure”: instead of arguing directly against the distributive law, Putnam brings up general relativity and non-Euclidean geometries and uses them to bring into question the very notion of self-evidence (this point is related to what Kripke calls “methodological ascent.” See fn. 50).

Here is the softening up: “Look, this has happened in geometry. Do you know anything about the relativity?” If you do, then you get into all the complicated arguments and are not sure whether they are right. [...] Then one will be very softened up. Then one is sort of invited to accept the moral: “don’t ask yourself how could it be that one of p and q is true and some other statement r is true, yet

‘p and r’ is definitely false and ‘q and r’ is definitely false. If you start thinking that way, you are adopting a bad habit in the same way if you were to start thinking how could it be that two lines can be perpendicular to the same line and yet converge to each other.” That is the example he [Putnam] gives. How can it possibly be? After all, necessary truths have turned out to be falsehood, to use his slogan. (Princeton Seminar)

This “procedure” has for Kripke little argumentative value; it is just a way of changing the subject and avoid confronting the question of how a logical principle as self-evident as the distributive law could turn out to be false. In his view, Putnam has not shown the relevance of the case of geometry to the case of logic, much less that we cannot trust our intuitions of self-evidence. His advice is uncharacteristically straightforward:

... don’t feel so softened up by these examples. If non-Euclidean geometry is really as unintelligible as Putnam thinks it is [...] it should be the task of the philosopher not to stamp his foot and say “But it is true!” – though no doubt it is. It should be for him to investigate what we can mean by saying it is true, and where we went wrong in our feelings of self-evidence.¹⁵⁷ At least this is a viable alternative position. One does not have the authority of Science with a capital S for the proposition that self-evidence is a category which must be given up. I mean, that may be true, but it certainly isn’t shown in any clear way by any example like this one. (Princeton Seminar)

And in the Pittsburgh lecture he says:

¹⁵⁷ In a seminar on Frege and in personal conversations Kripke has expressed sympathy for the view that there is still something to the traditional claim that Euclidean geometry is *a priori*. In the Princeton Seminar he says that Putnam doesn’t pay enough respect to the feeling of self-evidence that accompanies the idea that two straight lines both perpendicular to another line cannot meet. In his opinion, there is something to this feeling that we should at least try to explain. A similar point is made by Frege in the passage from *The Foundations of Arithmetic* quoted before (see fn. 151). See also Katz (1998). He argues that Putnam’s case “concerns an empirical application of Euclidean geometry, the case is one in which an *a posteriori* applied geometry was falsified on empirical grounds, not one in which an *a priori* pure geometry was falsified on such grounds.” (p. 49)

The fact that something can seem to be self-evident and later turn out to be wrong shouldn't undermine our using this as a principle of seeing that things are true, any more than the fact that something that seems to be supported by experiments and then later turn out not to be so well supported by experiments, should undermine our using 'being supported by experiments' as a justification for accepting something. (Pittsburgh Lecture)

That something *seemed* to us to be a self-evident truth and then turned out to be false does not show, according to Kripke, that we have to discard self-evidence as "a principle of seeing that things are true." Here, again, we face the question of what is it exactly that we see to be true: is it the UI principle, or that in a *particular* inference 'Fa' follows from $\forall x(Fx)$? In the previous passage he implies that the distributive law is self-evident, and he does say this law is "pretty evident" to him, but he doesn't say how this intuition of self-evidence connects with our intuitive reasoning according to it.

Two other points should be highlighted in connection to these passages. First, the last passage suggests that an intuition of self-evidence not only explains our taking a basic logical principle as true (or a rule as valid) – or, in the other interpretation, our taking a conclusion to follow from certain premise(s) in the case of a particular inference – but also justifies us in doing so. Second, though there is no particular problem in thinking that we can be wrong about something we considered self-evident, Kripke thinks that intuitive arguments are needed to show why we were wrong; a straight rejection or modification won't do. We have to dispel their air of self-evidence; for him, intuitions can only be contested with more intuitions, not with methodological arguments against self-evidence coming from other areas. What bothers Kripke about Putnam's defense of quantum logic is that he doesn't build his case by providing intuitive arguments against the distributive law; instead, he concentrates his efforts in the "softening up procedure." As a result, the distributive law is, in Kripke's view, as intuitive as it was before, but we are told

that it should be given up because the notion of self-evidence has been shown to be inadequate in any case, and the revision would allow a realist interpretation of quantum mechanics.¹⁵⁸

... there are two possibilities for changing one's mind about an alleged self-evident truth. One is, you see that though it seemed to be self-evident, you had overlooked something, and once you understand this, you really understand why the apparent feeling of self-evidence was just based on something you overlooked. The other way – and this has happened – is that you stamp your foot and say, “Look, it's false anyway,” “for empirical reasons!” or something. Like you say, “I still can't see what's wrong with why two plus two isn't four, or two times two isn't four. That seems just as obvious as it was before, but I get such a nice system of physics without it, that I will conclude that two times two is not always four.” It's not that I see now what I overlooked in applying it to quarks, or something like that, but we just stamp our foot and say that this does not hold in the quark theory.

Intuitive reasoning plays an important role in the two cases Kripke is willing to recognize – in a (very) loose sense – as “changes of logic.” The first one is the case we were considering: the rejection of what we thought was a self-evident principle. The “change” is the result of noticing either an overlooked case, or a mistake or hidden fallacy by means of our intuitive or usual way of reasoning. Kripke gives as an example the case of Aristotelian logic:¹⁵⁹

Once logicians have actually formulated what principles hold in logic, they *can* sometimes turn out to be wrong. They're wrong because of an intuitive argument, not because a different formal system was proposed and then adopted. The most famous case, and the one I think least open to dispute, is the case of the traditional logic, the Aristotelian logic, which accepted the inference from ‘All As are Bs’ to

¹⁵⁸ In another passage Kripke says: “Suppose it were the case that one could see what the mistake was. Then there would be no problem in giving it up. But suppose (this can be the only case with pizazz) you *can't* see what's wrong. It still seems to you just as self-evident as before, but you haven't found something you were overlooking or some hidden assumption, so that it seemed to be self-evident but it wasn't, but we still give it up because we have an alternative geometry ... We have an alternative which yields rich consequences and so on and so forth. And we adopt it, unintelligible as it may be to us, and intelligible as the alternative may be to us. That's the kind of case Putnam is considering.” (Princeton Seminar)

¹⁵⁹ He also mentions Einstein's critique of the notion of simultaneity as an example of a non-logical case where something that seemed self-evident was rejected by means intuitive arguments.

‘Some As are Bs’, and didn’t realize that this leads one into trouble if one allows empty terms. That is to say, a law of logic was accepted for many, many years, then it was later seen to be a mistake – not because someone said ‘Let’s adopt a different formal system’ (whatever that means) but because we know this is a mistake. A mistake was noted, namely, that we overlooked the possibility that the term ‘A’ is empty.

[...] there you have a famous case at least where something which used to be universally accepted is now widely disputed or widely thought to be false. This was not because of a Putnam kind of thing, but because of an intuitive thing. It wasn’t that they said, ‘That was one logic. Let’s adopt another.’ It was that this principle of inference, which used to be universally accepted, we now argue, giving an intuitive argument – that is, an argument independent of adopting a formal system – is, in fact, fallacious.

[...] in a complicated case one can certainly be wrong about what’s a law of logic. But even in a case where it seems self-evident, one can turn out to be wrong later by further intuitive reasoning. As in the case of the Aristotelian logic: a simple case was overlooked there, namely, the case where the subject term is empty in ‘All As are Bs.’ These cases shouldn’t be controversial, they shouldn’t undermine the notion of self-evidence. (Pittsburgh Seminar)

This case, Kripke says, may be considered a “change of logic” if we wanted to use such a “pompous phrase,” but it shouldn’t be seen as the rejection of a formal system and the adoption of another, or as undermining the notion of self-evidence.¹⁶⁰ It was motivated by intuitive considerations that lead us to see what was wrong with it and does not pose any major philosophical problem.

The second case of “change of logic” involves the introduction of new connectives which are given a different interpretation than the classical connectives; the non-standard laws or rules satisfied by the new connectives are specified by reasoning intuitively about the intended interpretation. Kripke thinks that this is the way intuitionistic logic should be viewed. Intuitionism is not, in Kripke’s opinion, committed to the rejection of the classical connectives as

¹⁶⁰ See Priest 2003: 454ff for a rather different reading of this case.

meaningless, it introduces different connectives which are supposed to apply to the mathematical case and receive a non-standard interpretation in terms of provability.¹⁶¹ The important point for our purposes is that intuitive reasoning is needed to figure out the logical laws or rules of inference that are supposed to capture the intended interpretation. Kripke says:

It's not that we set up the different laws and then, as the literature often says, "deduce" that they must have a different interpretation – so Carnap, and sometimes Quine, would argue. But rather, it's the other way around: One sets up a different interpretation and then sees that in virtue of this non-standard interpretation, non-standard laws follow. (Pittsburgh lecture)

Neither of these cases would, strictly speaking, be considered a "change of logic." But what matters to us, at any rate, is that in both cases intuitive reasoning (the *logica utens*) clearly takes priority over the logical principles or rules of inference (the *logica docens*): it is needed to reject alleged self-evident logical principles and to figure out the laws that capture a non-standard interpretation.

5.2 The Role of Intuitions and the Adoption Problem

In the last section we have tried to figure out how Kripke understands the connection between "intuitive reasoning" and self-evidence, but different passages appear to suggest different interpretations. It is, indeed, quite possible that Kripke did not have a clear view on this matter. Moreover, he doesn't try to say what this intuition of self-evidence or "seeing that something is self-evident" is supposed to be. He seems to have in mind something along the lines of what Bonjour calls "the traditional view" of rational intuition, according to which the intuitions result

¹⁶¹ See Dummett (1973a and 1991) for an opposing view.

from “direct or immediate insight into the truth, indeed the necessary truth, of the relevant claim.” (2005)¹⁶² Although in one passage Kripke mentions an “intuitive apparatus,” it is unlikely he means there is a specific faculty that is the source of these intuitions; he probably thinks that they are just the result of a capacity or ability we possess, perhaps simply in virtue of being rational.¹⁶³

An important point to mention is that Kripke is not particularly concerned with rational intuition as a source of epistemic justification. His interest in intuitions rests on the need to uncover something that motivates our reasoning in accordance with basic principles. The idea seems to be that once we grasp the truth of a certain basic logical principle, we ‘see’ what follows, and this somehow guides our “intuitive reasoning” in accordance with the principle in question. Of course, the propositional knowledge we acquire by means of such rational intuition will be *a priori*,¹⁶⁴ but that plays only at a second stage. The role of intuition is of interest primarily as the *source* of our reasoning, and although it would make sense to link the justification we may provide of such reasoning with its intuitive source, this is more a consequence of the view than a rationale for it.¹⁶⁵

¹⁶² Self-evidence does not, however, commit us to incorrigibility. As we have seen, Kripke makes room for the possibility that we might be mistaken in our feelings of self-evidence.

¹⁶³ A specific faculty for rational intuitions need not be presupposed. In fact, Bonjour, for example, denies that there is some distinct psychological faculty involved in attaining *a priori* knowledge. It is simply the “ability to understand and think” [see Bonjour (1998), p. 109]. And, as I said above, I don’t think Kripke is committed to a faculty of rational intuition.

¹⁶⁴ Kripke doesn’t explicitly state so, but it is safe to say that it follows from his talk of self-evidence and ‘seeing’ that a principle is true.

¹⁶⁵ Although I may be putting words in his mouth, since Kripke’s remarks on the role of intuitions are very unsystematic and his intention, most likely, was not to adhere to any particular brand of rationalism, I cannot help but notice some similarity between his views and Bonjour’s moderate rationalism. (See Bonjour 1998: chapter 4.) According to Bonjour we have a capacity for ‘rational insight,’ which is basically an intellectual act by which we are able to “see or grasp or apprehend in a seemingly direct and unmediated way that the claim in question cannot fail to be true.” (p. 101) Among the candidate claims are, for example, the proposition that nothing can be red all over or green all over at the same time, the proposition that two plus three equals five and, of course, logical principles. When the necessity of a claim

As noted in chapter 1, rational intuition views have been around for an extended period of time. The main criticism they face also has a long tradition: those that favor it have consistently failed to give a satisfactory explanation of what it is supposed to be and how it operates. Critics of the view often point out the lack of progress in this respect. The point can hardly be denied, but what's interesting is that *pace* these problems, it continues to draw strong support. And indeed the view seems to have an irresistible appeal when it comes to basic cases, like bachelors or logical principles.

We will, in any case, assume in what follows that we have a reasonable account of what these intuitions are and how they operate; the idea is to discuss the view from the perspective of the adoption problem. Since it is not clear what view, if any, Kripke holds, we will consider different possibilities. Two main interpretations of Kripke's passages were introduced in the last section:

- (1) It is because we find the UI rule self-evident that we find particular UI inferences evident as well and we reason "intuitively" according to them.
- (2) it is not the principles or rules what we see as self-evident, but the particular inferences themselves. According to this, being empty headed or unable to understand the meaning of, say, 'all' is simply being unable to *see* that "this is green" follows from "Everything is green," as Kripke suggests in one of the passages quoted before.

or proposition is apprehended by rational insight, it is thus considered rationally self-evident, which essentially means that "its very content provides, for one that grasps it properly, an immediately accessible reason for thinking that it is true." (p. 102) And, like Kripke, he rejects the view that such *a priori* insight is infallible. However, Bonjour conceives rational insight as a "fundamental source of epistemic justification" (p. 100), and claims that a direct insight into the necessity of a proposition or claim in principle justifies our accepting them as true. But this is not, as I argued above, Kripke's main concern.

Let's go back to Harry and assume that by an act of rational intuition he sees the truth of 'All universal statements imply all of their instances,' or that MPP is a valid rule of inference. Would that help him? It should be clear by now that the answer is 'no.' If telling him these things didn't help him, why would seeing that UI is self-evident be any different? Though the source of our coming to believe that UI is true is different, both come down to the same thing: the acceptance of the truth of a principle or the validity of an inference. But as we know, mere acceptance won't get us very far.¹⁶⁶

We can't thus expect someone like Harry to achieve much by holding UI to be self-evident, determining how the self-evident principle or rule is to be applied presupposes inferring according to UI. Wittgenstein seems to have this in mind when he says: "If intuition is an inner voice—how do I know *how* I am to obey it? And how do I know that it doesn't mislead me? For if it can guide me right, it can also guide me wrong. (Intuition an unnecessary shuffle.)" (*PI* § 213)

Similar arguments against rational intuition as a source of epistemic justification are offered by Boghossian (he has, essentially, Bonjour's moderate rationalism in mind). He argues (2003: 271ff) that rational insight into the validity of MPP (which he expresses as: for all p, q : necessarily: if both p and ' $p \rightarrow q$ ', then q) will not justify any particular MPP inference made by a subject.

Rational insight, we are conceding, gets us as far as the general propositional knowledge that all arguments of MPP form are valid. However, to bring this

¹⁶⁶ Once again, the problem is that when confronted with a particular case, let's say inferring 'Alex the lion talks' from 'all the animals in the movie *Madagascar* talk,' he may intuitively see the truth of the general principle, but to apply it to this case the intuition won't be sufficient, he will have to use it. To use it, however, he will not only need to perform an application of UI to derive the conclusion that 'Alex the lion talks,' but also from the intuited general principle to the particular case 'all the animals in the movie *Madagascar* talk,' since this case is an instance of the general principle. Therefore, to put the intuition to work, we need to apply the principle to the intuited principle itself. Harry, as you may recall, has never used UI before, the intuition was supposed to help him out, but now it turns out that to be able to put the intuition to good use, he needs to apply UI to the intuition. This is hardly any progress.

knowledge to bear on the justifiability of any particular inference will require the thinker to be able justifiably to infer the validity of that particular inference from the validity of all arguments of MPP form. And this will require him to be able to reason according to MPP justifiably. (Boghossian 2003: 273-274.)

A circularity problem arises here: the particular MPP inference could only be deemed justified for a certain subject if she is already justified in inferring according to MPP. In Boghossian's words, the problem is that "the very ability we are trying to explicate is presupposed by the internalist account on offer." (p. 274)

Maybe what we are missing here is that the rational intuition not only works at the general level, telling us that 'All universal statements imply all of their instances' is true, but also at the particular level, with each case. In the case under consideration then, we would get an intuition telling us: 'Alex the lion talks' is true, it follows! What we have here is not an intuition on the validity of each specific instance (this possibility is considered below), but a direct intuition about the truth of the conclusion. However, it seems to me that it would be very implausible that an intuition is required for each case. The main problem, leaving implausibility aside, is that if we need an intuition whispering in our ears what follows in each case, it is hard to see what would connect all those particular cases with the general logical principle that was supposed to be self-evident. The point remains that, so as to see each case as an instance of the general principle, the use of UI on the general principle itself is needed. Without that, the generality of the logical principles would be lost.

Boghossian also considers the possibility that we grant thinkers "rational insight into the validity of *specific* inferences, and not require that this be derived from some general knowledge of the validity of all inferences of the form MPP" (p. 274, his emphasis) as a way of avoiding the circularity problem. But he argues that this strategy will not succeed. The problem here, he

claims, is that a thinker's justification of a specific MPP inference rests on the thinker's justified belief that the premises of the inference give a good reason for the conclusion, but applying this to the specific case involves inferring in accordance with MPP. Boghossian unpacks the inference as follows:

We have a rational insight about the validity of a specific MPP inference, so we claim:

(1) This particular inference from (1) and (2) to (3) is valid.

But the justification of such inference depends on our justified belief that:

(2) If an inference is valid, then anyone who is justified in believing its premises and knows of its validity is justified in inferring its conclusion.

Applying this to the case at hand we get:

(3) Anyone who is justified in believing the premises of *this inference* is justified in inferring its conclusion.

(4) I am justified in believing the premises (1) and (2).

Therefore,

(5) I am justified in inferring (3)

Thus, the circularity has not, after all, been avoided: even if we are dealing with rational insights on the validity of specific inferences, MPP is needed (see Boghossian 2003: 274).

BonJour and Boghossian are both worried about the justification of logical principles, which, as I have stressed, does not seem to be Kripke's main concern. But Boghossian's arguments closely resemble the ones I gave against Kripke's views on rational intuition. Notice that a similar argument can be provided for UI. Contrary to the case considered above, where the intuition directly delivers the conclusion of the argument, we have here a rational intuition on the validity

of the specific UI inference, we grasp the validity of ‘all the animals in the movie *Madagascar* talk, then Alex the lion talks.’ But in this case we will also need to be in possession of a more general principle stating something like “valid inferences license the detachment of their conclusions,” to infer that ‘Alex the lion talks.’ And, needless to say, to apply this principle to the specific case we will need to apply UI, something that Harry just doesn’t know how to do.

It seems to me that the moral to be drawn is that recognizing the truth or validity of a certain principle or rule via rational intuition does not help a subject unless that subject was already capable of inferences of the relevant kind. The mistake here is to think that the intuition provides us with a standpoint that is somehow more basic than using the rule or principle itself. Presupposing an intuition not only does not explain or justify our use of basic logical principles, it makes quite impossible to see how those uses could ‘spring’ from the intuitions. Intuitions lying beneath the use seem to be inert, without effect. It would be in principle possible to actually have those intuitions and never use them. As in the other cases, we need to be granted the ability to reason to be able to do something with them. Thus, it seems that the Kripke-Carroll argument not only goes against views on the empiricist tradition. It also goes against views that favor the *a priori* and are usually considered part of the rationalist tradition. And it also goes against Kripke’s own views on the role of intuitions.

Wright has argued that views based on rational intuition may be guilty of ‘empty self-congratulation,’ because no one has been capable of explaining how rational intuition is supposed to work and be integrated with our scientific knowledge. Wright no doubt has a point there, but what I have been trying to show above is that there won’t be anything to congratulate ourselves about. At this point, we may want to recall Wittgenstein’s remark about intuitionism. In his distinctively blunt style he says: “Intuitionism [...] requires that we have an intuition at each

step in calculation, at each application of a rule [...]. We might as well say that we need, not an intuition at each step, but a *decision*. –Actually there is neither. You don’t make a decision: you simply do a certain thing. It is a question of a certain practice. Intuitionism is all bosh – entirely.” (Wittgenstein 1989: 237) And in *On Certainty*, he says: “Giving grounds, however, justifying the evidence, comes to an end; - but the end is not certain propositions’ striking us immediately as true, i.e. it is not a kind of *seeing* on our part; it is our *acting*, which lies at the bottom of the language-game.” (§ 204)

Chapter 6

Concept-Constituting Accounts

We remarked several times that a natural reaction to the adoption problem would be to say that neither Harry nor Ernie understand what we mean by ‘all.’ The fact that Harry does not infer in accordance with UI and he is still unable to do it after we tell him the relevant rule or principle would then indicate that there is something wrong with his understanding of ‘all,’ that he is just not getting what that means.¹⁶⁷ Analogously, Ernie’s inferring “Alex the lion is not vegetarian” from “All lions are vegetarians” would indicate that he either does not comprehend the meaning of ‘all’ or has a different understanding of it. (Especially when we take into account that these are not isolated cases).

Nothing particularly special would then appear to be going on here: similar thoughts are to be expected if, for example, we happen to witness the unlikely denial of, let’s say, “Bachelors are unmarried men.” For, what is there to doubt or deny? Assuming that we need not worry about anything else unusual going on, the thought that whoever denies this statement is not getting the meaning of one or more of the words contained in it seems inescapable.

Such intuitive reactions suggest a connection between the way words or concepts are used and our understanding of them. And, of course, that there has to be some sort of connection is likely to strike us as a platitude. The problem is how to account for this exactly. A popular way of cashing out this connection is to make it constitutive. Such is the route followed by some versions

¹⁶⁷ Or, as Michael Levin has remarked to me, that he does not understand ‘instance.’ The point is that it is hard to see how Harry could understand both ‘all’ and ‘instance,’ accept “all universal statements imply each of their instances” or “From any universal statement, you may deduce any instance,” and at the same time not conclude ‘*Fa*’ from ‘*VxFx*.’

of conceptual role semantics:¹⁶⁸ certain inferential transitions are concept (or meaning) constituting, and it is a condition for our actually possessing the concepts contained in them (or understanding the meanings) that such transitions are somehow endorsed (more on what kind of endorsement could be appropriate below). There is a “traditional connection” (Wright and Hale: 2000) between this semantical view and the epistemological notion of justification: the justification for the constituting inferential roles is secured by our possessing the relevant concepts. And, given that their justification need not depend on experiences of the external world, their justification is *a priori*. They are, indeed, analytic.

Concept constituting accounts have enjoyed enthusiastic support for some time now, with a number of influential writers developing different versions of this account. Although the idea that some sentences are analytic is not precisely new, Carnap’s attempt to account for logical truths via conventionally stipulated implicit definitions is an important predecessor of the concept constituting accounts. As we know, Quine played a key role in discrediting the view. The success of his criticisms is shown by the fact that the discredit went beyond the specifics of Carnap’s proposal, covering the notion of implicit definition itself and instilling suspicion into any attempt to earn justification by way of implicit definitions.

Those that find concept-constituting accounts attractive, however, argue that the suspicion has gone too far. Quine’s criticisms to Carnap’s proposal are mostly accepted, but they have attempted to rescue the “traditional connection” between conceptual role semantics and *a priori* justification from the Quineans by dropping its association with conventionalism. Though analyticity as a conventional stipulation of the truth of an assertion cannot be saved, there

¹⁶⁸ A family of related views goes under the name of conceptual role semantics. See Block (2000), and Greenberg and Harman (2007: 242-256).

remains the possibility, in their view, of understanding it as conferring justification, provided the stipulations are needed to constitute one of the concepts contained in them.¹⁶⁹

Although concept-constituting accounts are thought to be particularly compelling when the meaning of the logical constants is at issue,¹⁷⁰ they are not only meant to be applied to logical words.¹⁷¹ Gentzen's suggestion that the introduction and elimination rules of his Natural Deduction System define the meaning of the logical constants provided the model for logic, which was later on generalized to other domains. Dummett, Martin-Löf, and Prawitz were among the first to exploit Gentzen's idea by developing different accounts of the meaning of the logical connectives in which understanding is given by the acceptance of the introduction and elimination rules.¹⁷² But it was Dummett who first thought that Gentzen's model could be

¹⁶⁹ In Boghossian's words: "The first and most important task was to detach the idea of epistemic analyticity – justification on the basis of meaning alone – from the idea of metaphysical analyticity – truth on the basis of meaning alone – and to show that the former idea could survive without any help from the latter." (2008: 4) For discussion of the differences between the projects, see, for example, Boghossian (1996, 2004), Peacocke (1993), Williamson (2007, ch. 3 and 4), and Wright and Hale (2000).

¹⁷⁰ Fodor and Lepore write: "We think it's a *prima facie* objection to this project that 'and' and (as it might be) 'tree' seem to be such different *kinds* of words that it would be sort of surprising if their meanings were constituted in essentially the same way. (For one thing, it's plausible that 'and' doesn't refer to anything; but 'tree' clearly does.) Given the *prima facie* magnitude of such differences, it seems CR semantics might do well enough for the logical expressions but not work at all for non-logical primitives. One of the present authors (JF) has flirted with this sort of hybrid theory, and continues to find it attractive on alternate Tuesdays." (2006). See also Fodor (1990: 110-111).

¹⁷¹ Peacocke says: "the idea of an implicit conception [what we have called, following Wright and Hale (2000), the traditional connection] underlying a concept is entirely general and can in principle be found in almost any domain, well beyond those of logic and mathematics." (2000: 283).

¹⁷² In fact, Gentzen suggested that the introduction rules, in particular, defined the logical constants. He says: "the introductions represent, as it were, the 'definitions' of the symbols concerned, and the eliminations are no more, in the final analysis, than the consequences of these definitions" (Gentzen 1932; see Koslow 1992, part 1, for discussion) The purpose of a natural deduction system was for Gentzen to "set up a formalism that reflects as accurately as possible the actual logical reasoning involved in mathematical proofs." A clear view of the interaction between the rules is important to avoid *tonk*-like connectives (see below). Dummett's account of the harmony requirement between the rules takes Gentzen's suggestion to heart. In contrast, Martin-Löf, for example, has chosen the elimination rules to do the bulk of the work.

applied, with appropriate modifications, to the language as a whole.¹⁷³ More recently, different versions of conceptual role semantics (also known as inferentialism) have been held by authors such as Brandom, Bealer, Peacocke, Hale and Wright, Wright, and Boghossian.¹⁷⁴

Coming back to our problem, we can hardly deny that there is something wrong with Harry's and Ernie's understanding of 'all.' In view of what we have been saying, a possible diagnosis of Harry's situation would be to claim that he lacks the concept 'all' (or 'instance') Assuming the concept 'all' is constituted by its introduction and elimination rules, Harry's inability to infer in accordance with UI can be taken to support such claim. For, if the UI rule partially constitutes 'all,' not being able to follow it would indicate that he does not possess the concept. Moreover, if having the concept is, as some versions of conceptual role semantics suggest, necessary for understanding the word,¹⁷⁵ Harry's lack of understanding should not be surprising at all. His inability to perform universal instantiations does not just *suggest* lack of understanding, it simply *entails* it (we can assume that his understanding of the non-logical words in the original example is

¹⁷³ Dummett thought that a proper account of the meaning of the connectives should stem from a general theory of meaning, which for him was also a theory of understanding. Throughout his work he insisted on the idea that the question for what general form a theory of meaning ought to take, whether in terms of truth-conditions or verification or provability conditions, was to be decided first. According to his "bottom up" strategy, once we have decided the form that the theory of meaning has to adopt and the role that the notion of truth has to assume in it, it would be possible to justify a notion of valid inference and a certain image of the metaphysical character of reality will impose itself on us. (See Dummett 1991:12). As is well known, he favored a theory of meaning that took the notions of verification and proof as central (instead of the notion of truth), an intuitionistic account of validity, and an anti-realist view on metaphysics.

¹⁷⁴ See, for example, Brandom (1994, 2000), Bealer (1999), Peacocke (1992, 1993, 2004, 2008), Hale and Wright (2000, 2001), Wright (1999), Boghossian (1996, 2000, 2001a, 2003, 2004).

¹⁷⁵ This would depend on whether concepts or meaning take priority, which in turn depends on whether thought takes priority over language, or conversely. More recent accounts tend to favor concepts, while for Dummett meanings were the undisputed central notion. He actually held that the idea that thought could take priority over language falls outside analytic philosophy, since in his view the priority of language its supposed to be one of its main theses. (See Dummett 1993. For a –in my view correct – rejection of the view that analytic philosophy can be characterized by a set of theses, see Williamson 2007.)

not in question). And analogous considerations can be provided for the case of the perverse instantiator.

Pace the initial plausibility of this diagnosis of Harry's troubles, I will argue below that concept-constituting accounts cannot easily explain away Harry-like cases, and that in fact the adoption problem brings to light a flaw in these accounts. In the remainder of this chapter, we will focus on Boghossian's version of such account. For reasons that will (hopefully) become clear as we move on to the next chapter, not only Boghossian's actual proposal but also the arguments which led him to it (whether the process of thought presented is Boghossian's actual mental history or just a possible one) and the objections that have more recently brought him to develop second thoughts about it, will be relevant for the issues we have been discussing in connection with the adoption problem.

6.1 Clearing the Ground for a Circular Justification

Is it possible to justify our logical beliefs? Boghossian's answer to this question is the result of a carefully crafted argument that moves forward by elimination, discussing and rejecting a wide range of proposals along the way. By the time he settles on his preferred account – i.e., a rule-circular justification of basic logical beliefs grounded on a concept-constituting account of warrant transfer by inference¹⁷⁶ – one is meant to think that he (and the reader as well) has been cornered into this view. Yet the rejected views are not simply there to build suspense. A justification that involves circularity is unlikely to be our first choice. Therefore the inadequacy of

¹⁷⁶ See Boghossian 2000 and 2001a for a defense of the whole project, and 2003 for his account of warrant transfer by inference.

the other views is essential to his argument, and Boghossian does indeed go to great lengths to present his proposal as our only hope.

Although we will be mainly interested here in Boghossian's positive account, it will be instructive to review his objections to other proposals. Views that deny the possibility of a substantive justification are rejected first. In chapter 1, we said that Boghossian opposes the claim that our fundamental logical beliefs cannot be justified. His main argument builds on the following principle:

(LP): We can be entitled to reason in accordance with a logical rule only if the belief that the rule is truth-preserving can be justified.

Stated contrapositively:

(LP'): If it is impossible for us to be justified in believing that a certain logical rule is truth-preserving, we cannot be entitled to reason in accordance with that rule.

(Boghossian 2000: 234)

The principle, which should be respected by "any plausible epistemology of logic," (p. 235) establishes a link between the *logica docens* and the *logica utens*: it is necessary for our *beliefs* in the fundamental logical rules to be justified if our entitlement to *use* MPP inferences is to be granted (though, as Boghossian stresses, the principle doesn't say whether beliefs or uses will be first in the order of explanation. See 2000: 234, fn. 7). He makes clear that what's at stake here is not just the truth-preservingness of our fundamental rules, but *our claim* that they are so, already

suggesting that he finds ‘externalist’ accounts of knowledge of basic logical beliefs inadequate (we will come back to this).¹⁷⁷

What’s wrong, then, with the claim that our logical beliefs are unjustifiable? Let’s assume with Boghossian that we are restricted to a propositional system where MPP (his favorite rule of inference) is the only underived rule. According to (LP), the consequence of contending that the belief that MPP is truth-preserving is unjustifiable is that we won’t be entitled to perform inferences that accord with the MPP rule. As a result, we won’t be able to claim to be justified in upholding the beliefs that were obtained by such inferences either. But then, Boghossian points out, if MPP is the only underived rule, any belief that was obtained using deduction would be unjustified as well. Moreover, *this* very argument is itself the product of deduction, and thus we can’t be justified in believing its conclusion either. Surely, as he observes, this is not a stable position to come to rest. (2000: 234-35)

It is difficult, once (LP) has been granted, to resist Boghossian’s claim that going down the skeptical path demands more work than just maintaining that our fundamental logical beliefs cannot be justified. Such efforts would in any case be ill-advised from Boghossian’s point of view, considering that he finds the different candidates for bringing balance to the skeptical outlook ultimately unable to do a credible job.

The first two candidates, relativism and non-factualism about logic itself, are quickly deemed hopeless. Relativism about logic is “just a version of a conventionalism about it,” (2001a: 247, fn.

¹⁷⁷ Wright has argued that externalist accounts of knowledge fail to address skeptical worries about our *claims* to knowledge. He says: “... what is put in doubt by sceptical argument is —of course— not our possession of any knowledge or justified belief—not if knowledgeability, or justification, are conceived as constituted in aspects of the external situation in which we come to a belief. (How indeed could armchair ruminations show anything about that?) What is put in doubt is rather our right to claim knowledge and justified belief. It is this which the project of making out entitlements tries to address and which, on what seems to me to be a correct assumption, externalism is impotent to address.” Wright (2004a: 210; see also 2004b).

13) and non-factualism is also subject to the objections raised by Quine against conventionalism. Boghossian thinks that the way of making sense of a non-factualist view of logical beliefs is to hold that we didn't have any logical beliefs to begin with. To make such a view clear, he proposes an expressivist construal of (apparent) sentences reporting logical beliefs in terms of sentences expressing our acceptance of a system of norms that licenses the implication in question. Following Quine, he proceeds to argue that, since these norms will have to be general, their application to particular inferences will presuppose further inferences not explained by the non-factualist account itself. (See 2000: 237-38.)

The next two alternatives to consider, relativism and non-factualism about the notion of justification, are no less problematic. On his view, a relativism or non-factualism about justification disputes either the objectivity or factuality of our epistemic norms for belief change. Given the restriction to MPP, the only deductive epistemic norm would be:

If S is justified in believing that p, and is justified in believing that 'if p, then q', and S infers q from those premises, then S is *prima facie* justified in believing q. (Boghossian 2001a: 236).¹⁷⁸

Now, a relativism about justification maintains that justification is bound to the epistemic practices of a community of thinkers: a thinker's justification for a given belief is relative to the epistemic principles and norms of the thinker's community. There is no higher tribunal to appeal to for judging the justification of the community as a whole. Hence, whether the MPP epistemic norm *itself* is justified or not does not amount to a genuine question. Boghossian's main objection

¹⁷⁸ A parallel argument can be given in terms of epistemic principles. Another reason for concern is that MPP will be involved in the justification of other epistemic principles, and a failure to justify it will carry on to other epistemic principles. This, Boghossian fears, will lead to a global epistemic relativism where any epistemic principle is as good as any other.

is that the notion of being ‘epistemically blameless,’ which for him is a central epistemological notion, cannot be properly explained by this version of relativism. According to him, for any thinker to be epistemically blameless with respect to a belief, it should not make sense for us to criticize her holding such belief – and Boghossian intends this notion to be absolute (i.e., not dependent on communal acceptance). (2001a: 250) In a relativist construal of justification, blameless belief is cheap (or, even better, it’s free): any background beliefs could be made consistent with our belief in any proposition, provided that the adequate changes are made in the community’s epistemic norms. And hence, in Boghossian’s view, it cannot escape falling into an extreme form of relativism whereby any belief is as good as any other.

One might harbor some doubts about this criticism, for why would the relativist allow for an absolute notion of blameless belief? It would seem that she has good reasons not to do so, since from her point of view this absolute notion may not even make sense. And if blameless belief is conceived as intra-communal, there might be a way of regulating the community’s beliefs and epistemic norms.

At any rate, we’ll assume that Boghossian is right in thinking that this form of relativism leads to an extreme form of relativism. Can we do any better with non-factualism about justification? Boghossian finds similar objections to this alternative. If non-factualism is construed in expressivist terms, assertions such as ‘x is a justified belief’ would merely express the “acceptance of a system of norms that permits that belief under those circumstances.” On his view, basic forms of disagreement (that is, disagreement over the epistemic norms that each accepts) cannot be coherently explained. The point here is that – intuitively – when we disagree about basic epistemic principles we are not just saying that we don’t accept them but that they should not be accepted, either by us or any others. Period. The non-factualist, on the other hand, has no right

to make such a demand. In Boghossian's opinion, the non-factualist lacks the resources to make sense of the disagreement itself: if there are no facts to ground our preference for certain epistemic principles and norms over others, any belief could be vindicated given the appropriate set of basic principles and norms, and the notions of a belief being blameless or blameworthy lose their force.

To this, the non-factualist may produce a quick reply: so what? She is unlikely to be concerned about losing an absolute notion of blameless belief. In fact, she has already accepted that there cannot be such thing. And even if she were to grant that non-factualism cannot make sense of basic disagreements, this doesn't seem a fatal objection to the view. In any case, it is never easy to make sense of basic disagreements, and some have held that when it comes to basic logical principles, disagreement is not even possible (see Quine 1970b, ch.6).

Disagreements apart, if Boghossian is right, maintaining that we cannot justify our logical beliefs is by itself incoherent. Furthermore, it cannot be supported by appealing to relativism or non-factualism about logic without leading to conventionalist-like objections, nor by appealing to relativism or non-factualism about justification without leading to an extreme form of relativism. None of this sounds, of course, like a good option, and once again we seem to arrive at the claim that "it is not open to us to regard our fundamental logical beliefs as unjustifiable." (2000: 253. See also 2001a: 265)

As soon as we come to terms with this, Boghossian suggests two alternatives for the justification of logical beliefs: one inferential and one non-inferential. We have already discussed what he calls the 'master difficulty' of the inferential alternative: some kind of circularity will be involved while arguing for the MPP belief. Non-inferential justifications don't come across as especially promising either. Some of Boghossian's arguments against rational intuition and

Quine's empiricist proposal were provided in chapters 4 and 5. Among the other candidates he considers, we find arguments based on the notion of conceivability and on the idea that fundamental logical beliefs are default reasonable. The impossibility of conceiving or imagining counterexamples to basic logical rules is not, according to him, legitimately non-inferential. The search for counterexamples cannot be done without the use of logical inferences (Boghossian 2001a: 231). He is somewhat more sympathetic to the idea that some basic beliefs might be "default reasonable" and, hence, reasonably believed without the need of supporting reasons. For him, the prospects of this view depend on whether we can figure out a principled way of determining which beliefs qualify for such status. Nevertheless, he doesn't envision how this could possibly be done (Boghossian 2000: 238-39).

Seeing as we are now running out of alternatives, Boghossian's subsequent suggestion that we were probably too hasty in discarding the inferential path is unlikely to strike us as unreasonable. His elimination strategy would be already paying off if we felt that way.

6.2 Vindicating Rule-Circularity

Before getting into Boghossian's argument, some distinctions and clarifications are in order. First, recall that Boghossian restricts his argument to propositional logic and a system that has MPP as its only inferential rule. Second, as LP illustrates, he distinguishes between beliefs and uses. We have dwelt on this distinction before. Here is how he draws it.

He distinguishes epistemic rules and inferential rules:

Epistemic Rule (ER): If you are justified in believing that p , and justified in believing that ‘if p , then q ’, then you should believe q or give up one of the other beliefs.

and

Modus Ponens Inferential Rule (MPPR): $p \rightarrow q, p / q$

from epistemic principles and logical implications:

Epistemic Principle (EP): If S is justified in believing that p , and is justified in believing that ‘if p , then q ’, and S infers q from those premises, then S is *prima facie* justified in believing q .

and

(MPP) $p, p \rightarrow q$ imply q . (Boghossian 2001a: 238-39)

How do these rules and principles relate to each other? Well, we follow the epistemic rule ER (and other non-deductive rules) to modify our beliefs. By doing so we exhibit, on the one hand, our acceptance of the Modus Ponens Inference Rule (MPPR).¹⁷⁹ ER only tells us, however, how we should employ it (it is “a rule of obligation governing belief”) but not that we *ought* to do so. Taking for granted that we seek to have justified beliefs because we aim for true beliefs, our reason for employing ER is that it gives us justified beliefs from justified premises. Thus, “in adopting [ER] as our rule of belief modification [...] we are implicitly committed to the truth of

¹⁷⁹ Boghossian claims that his argument is not affected by how the relationship between these rules is exactly understood, that is, by which rules count as more basic.

the epistemic principle.” (2001a: 236) Finally, a logical fact, that is, the logical implication MPP, has to obtain for EP to be true.

So, in their interrelations among each other the principles and rules remain true to the (LP) principle: for us to be able to know that we are justified in believing that we ought to follow the rule ER, it must be possible to know the principle EP, which in turn requires us to be justified in believing the implication MPP.

Now, what does Boghossian mean exactly by saying that we *follow* the rule ER? He is somewhat hesitant over this issue:

By saying that a thinker has and operates according to these rules I don't mean that the thinker grasps these rules as propositions. I mean that he follows these rules, and that this shows up in his behavior, however exactly that is to be analyzed. It will do no harm, for present purposes, to think of rule-following as a disposition to rule-conform under appropriately idealized circumstances. (2001a: 236)

According to this, no propositional grasp of the ER rule is required to follow it. In a footnote, however, he refers to the discussion in Boghossian (1989) and (2008). If we look at these references, we find that he claims there that “to say that S is following rule R is not the same as saying that S’s behavior happens to *conform* to R. Conforming to R is neither necessary nor sufficient for following R” (2008: 116). For us to follow ER, our acceptance or internalization of the rule has to be our reason for drawing a particular conclusion, it must “explain and rationalize why we form the beliefs that we form.” (119) This issue will be important later on, but let’s leave it aside at the moment.

Reconsidering the inferential path requires from us to deal with the “master difficulty” we have encountered several times before: any inferential attempt to justify basic logical principles

will be subject to circularity objections. Boghossian's first step in addressing this question is to distinguish between grossly circular arguments and rule-circular arguments. Whereas a grossly circular argument for MPP would simply include MPP as a premise, a rule-circular argument that, for instance, appeals to the material conditional's truth-table would make, as we have seen, one inferential step according to MPP rule.

The reasons for rejecting grossly circular arguments are obvious: as a justification for MPP, it would be worthless. The argument doesn't prove anything, it simply begs the question by assuming what we intend to justify. In addition, there is a further problem that Wright and Boghossian have labeled "bad company" (2001a: 243ff): anything can be 'proved' by such means, if there are no restrictions as to what we may wish to assume, and assuming is all that 'proving' requires.

Rule-circular arguments, though commonly used to explain why basic inferences are valid, have hardly been considered a serious option for justifying basic logical beliefs. A justification for the MPP associated conditional that makes an inferential appeal to the MPP rule strikes us as a bit of a cheat, a way of begging the question. How can one trust such a justification when it relies on an unjustified application of the MPP rule?

Boghossian contends that, contrary to appearances, it is possible for rule-circular arguments to have a justificatory role – provided certain restrictions are kept and certain expectations lowered. For his proposal to work, however, rule circular arguments will have to be rescued from the threats of begging the question and bad company.

The bad company problem looks less severe for rule-circular arguments. Boghossian stresses that these arguments might not imply that a rule is truth preserving. The main problem here is that there are well-known reasons for thinking that not only truth-preserving rules will go

through. One such reason is provided by Prior's famous *tonk* connective. Prior (1960) argues that it is possible to define a connective by means of the introduction and elimination rules, each rule being acceptable in itself, and show that the rules allow the derivation of any statement from any other in the following way:

- (1) For any statement P , we can infer a statement that contains the connective *tonk* and any other statement Q , that is, $P \vdash P\text{-}tonk\text{-}Q$.
- (2) For any statement $P\text{-}tonk\text{-}Q$, we can infer the statement Q , that is, $P\text{-}tonk\text{-}Q \vdash Q$.

Inconsistency arises immediately, given that these rules allow us to infer any statement Q from any statement P , licensing the inference $P \vdash Q$ as a derived rule of inference.¹⁸⁰

Boghossian reproduces a derivation due to Wright (2001: 245-246) where a rule circular argument beginning with a meaning postulate for the introduction of 'tonk' and making one inferential step with the 'tonk' rule is used to justify the associated conditional of 'tonk-introduction.' Wright (2001) emphasizes that similar derivations would be possible for any rule, and provides a derivation of denying the antecedent starting with a meaning postulate for the conditional. The moral here is clear: there better be a restriction to filter these cases if rule circular arguments are to be of any interest.

¹⁸⁰ Belnap (1962) specifies conditions that a set of logical laws has to meet in order to determine the meaning of a constant and avoid cases such as '*tonk*.' One such requirement is that the connective has to be a *conservative extension* of the language: there must be a way of obtaining the conclusion of the elimination rule from the premises of the introduction rule that does not appeal to the rules governing that constant. Dummett's harmony requirement is also meant to avoid such cases. (See Dummett 1991)

As to the other problem, Boghossian claims that it is not obvious that rule-circular arguments beg the question.¹⁸¹ To see why, we need to have a better grip on whose question we would be begging, and what exactly we expect such an argument to achieve. He readily admits that a rule-circular argument won't help dispel skeptical doubts about the validity of MPP. Consider the skeptic we encountered at the party in chapter 1. He doubted the validity of an inference we had made, and our attempt at appealing to the material conditional's truth-table to establish the validity of MPP inferences left him quite unimpressed. His attitude – though maybe provoked by a doubt that seems ridiculous in the given context – is in general perfectly understandable: if we doubt a claim A, we are not going to be persuaded by any argument which relies on A in its way to establish A. Clearly, such an argument is not going to satisfy anyone who had initial doubts about the validity of MPP.

Boghossian is unwilling to let a skeptic discourage him. He doesn't think that skeptical doubts are a good enough reason for discarding rule-circular arguments. For, why should we be concerned with the anyway probably impossible task of keeping the skeptic happy? We should not, in his view, let the skeptic decide what amounts to a legitimate reason for believing something. If we can come up with an argument that gives *us* reason to believe that the MPP rule is necessarily truth-preserving, the fact that such an argument won't get us any closer to persuading the skeptic should not be a reason for abandoning it.

If we are to believe Boghossian's elimination strategy, it doesn't look anyway as if we have much of a choice: we can't accept that no justification is possible without risking severe relativism, and the other alternatives seem to either be concealing their inferential nature or

¹⁸¹ Dummett thought that we needed to distinguish between 'suasive' arguments, those that intend to *persuade* someone of its conclusion, and 'explanatory' arguments, those that intend to *explain* the conclusion to someone who already accepts it. He argues that in the case of the justification of deduction only the latter arguments are available. See Dummett (1973: 296); also Dummett (1991: 202).

taking the justification of logical beliefs as a primitive. Throwing the skeptic under the bus seems to be the most reasonable choice.

As attractive as getting rid of the skeptic might seem, we shouldn't ignore that the skeptic's reaction to a rule-circular argument is not alien to *us* – that is, people that (I take) do not doubt the validity of MPP. The skeptic is voicing a demand that seemed plausible to us and that we had hoped to fulfill. Witness to this is that a discovery of rule-circularity in validity claims is usually viewed as disappointing or unsatisfying, and as a reason for looking for firmer grounds elsewhere. It should be clear that, in the context of a rule-circular argument, something else needs to be said – Boghossian explicitly acknowledges this point – about why it is okay to use MPP in an inferential step without having earned a justification for its validity.

6.3 Warrant Transfer by Inference and Concept-Constitution

Now, if we are going to put our trust in a rule-circular justification, we have to find a way of showing how we are entitled to the *use* of the rule needed for deriving the justification without having a prior justification of the *belief* that the rule is truth-preserving (or, more precisely, that the associated conditional is true or necessarily true). What we are after, then, is an account of the transmission of warrant from the premise(s) of an inference to the conclusion capable of holding its ground independently of any validity or truth-preservation claims. If we manage to secure such a warrant, then we can go back and seek a justification for the belief itself by rule-circular means.

As we anticipated, conceptual role semantics is crucial to Boghossian's understanding of the transmission of warrant by deductive inference. He intends his own account to address the

problems of two unacceptable extremes: an externalist view of warrant-transmission on the one hand, and an internalist view on the other. He claims, however, that though his proposal emerges as a sort of middle ground between the two, it is meant to be in line with an internalist account of warrant transmission.

For externalism – the position Boghossian labels “Simple Inferential Externalism” – the emphasis is placed on the validity of the inference: assuming the thinker’s reasons for believing the premises do not depend on a prior belief in the conclusion, the validity of the inference in question is all that is needed for warrant to be transmitted. The main problem with this view is that warrant-transmission comes too cheaply. Boghossian argues that an externalist account doesn’t provide intuitive conditions, seeing as many inferences are truth-preserving and still too complex to assume that a thinker would be warranted in holding the conclusion. He gives Fermat’s last theorem as an example: we wouldn’t intuitively think that a thinker that infers an instance of (1) from a particular claim of the form (2) is justified in believing (1):

1. $x^n + y^n \neq z^n$
2. x, y, z and n are whole numbers and $n > 2$

Yet this inference is necessarily truth-preserving – and, needless to say, analogous examples are not difficult to come up with.¹⁸² The point is that by making the mere reliability of truth-preserving inferences the central condition for warrant-transmission, externalism allows the inferring thinker to be completely ignorant of the status of the reliable process. The thinker is more of a medium, as it were, for the reliable process to carry on than an active and inextricable part of it.

¹⁸² The counterexample is analogous to Bonjour’s reliable clairvoyant Norman and other similar counterexamples where reliability seems insufficient to provide an intuitive account of justification. Boghossian views his objection as a specific application of a general problem for reliabilism. See Bonjour 1985: 37-41.

We saw before that, quite independently of the specific case of acquiring warrant by inference, being epistemically blameless is for Boghossian a key epistemological notion. More precisely, it is an intuitive requirement for justification: if a thinker is justified in holding a belief, it must not be possible to criticize her for holding it. The example of Fermat's last theorem indicates, however, that by asking too little of a thinker, externalism allows for epistemic irresponsibility. So it fails, in Boghossian's view, to provide sufficient conditions for the transmission of warrant. (See Boghossian 2001a: 257ff and 2003: 268ff)^{183 184}

An obvious way to avoid epistemic irresponsibility is to make the transfer of warrant more demanding for the inferring thinker. If the validity of the inference were known to her, charges of irresponsibility could not come so easily. However, a more modest and plausible condition, one that doesn't ask the thinker to actually know the validity of the pattern of inference, but only that it is possible to know that the inference is valid, may perhaps be enough to avoid charges of irresponsibility. Boghossian's "Simple Inferential Internalism" follows this idea by requiring the validity of the inference to be "reflectively accessible" to the thinker. Although such a requirement would make the proposal incompatible with a rule-circular justification, the previous considerations on epistemic irresponsibility would appear to lead us to it.

Boghossian can hardly be blamed for not seeing much reason for hope in this version of internalism. How is the validity of the inference supposed to be "reflectively accessible" to the

¹⁸³ Boghossian argues that a restriction to 'simple' inferences to avoid this kind of counter-examples would not help an externalist account. The reason is that externalism will have trouble explaining what 'simple' means in this context: if 'simple' means that the validity of the inference can be easily taken in, it is the thinker's capacity to take in the validity of the inference and not the reliability of the process what is doing the relevant work. If 'simple' doesn't involve taking in the validity of the inference, the problem is that there is no obvious basis for such a restriction. It's unclear how the relevance of a simplicity constraint for justification could be adequately explained. (See 2003: 269)

¹⁸⁴ See Rey (1993, 1998) for a defense of a reliabilist account of *a priori* knowledge of logic. In Rey's account logical competence is generated by a natural deductive system that is sub-conceptual and wired in the brain.

thinker? Something will have to be said about that, and at this point we have all grown weary of such attempts. As he notes, inferential arguments are out of the question at this level – and this holds for both, those that own up to their inferential pedigree, and those that don't.¹⁸⁵ A non-inferential account that relies on rational intuition, on the other hand, fails (as we have already seen in the last chapter) not only at the level of giving a justification for the belief that, say, MPP is valid, but also at the level of the inferential transition itself. Thus it is hard to tell what could work here.

In any case, Boghossian's main argument against Simple Inferential Internalism is that the accessibility condition is intuitively too demanding. One reason is that children (and maybe even some animals) cannot plausibly be required to have views about validity or about the rules involved in their inferential acts. Recall the case of Matías, my nephew, who is perfectly able to perform inferences that conform to MPP and UI, though he has never heard of validity and has no idea that his inferences can be seen as applications of some rules. We certainly cannot assume that such notions would be accessible to small children (much less to animals).

The other, more important, reason is connected with Boghossian's reading of Carroll's note. An important moral of the note is, on his interpretation, that assuming that for warrant to be transmitted a thinker must be able to know that the inference is valid leads to circularity or regress. The argument has already been rehearsed in our discussion of rational intuition (chapter 5). What should be emphasized here is that Boghossian understands Carroll as maintaining that an internalist construal is too demanding for warrant to be transmitted by inference at all.

¹⁸⁵ Such knowledge cannot be explained by inferential means since what is at issue is not the justification of logical beliefs but rather our entitlement to basic rules, which would be presupposed by any inferential account.

Let's take stock. On an externalist construal, warrant-transmission is cheap, it confers warrant by asking too little of the thinker. Internalism, in contrast, errs by asking too much, it makes the transmission of warrant unattainable by being too demanding. The way out of this is suggested to Boghossian by Wittgenstein's insistence on blindness:¹⁸⁶ the lesson to be drawn from the externalist's epistemic irresponsibility and the impossible demands of this version of internalism is, Boghossian claims, that a proper account of warrant-transmission should allow the thinker to be "blind yet blameless." His proposal intends to be a sensible form of internalism, where a thinker is, on the one hand, allowed to be blind – that is, warrants can be acquired without the need to be in position to be able to know that the inference is valid – and yet, on the other hand, blameless, not open to criticism regarding the beliefs that result from the inferential transitions.

It isn't clear that being blameless (or not being epistemically irresponsible, if they are the same) suffices for the reconstruction of an internalist warrant, and both Wright (2001) and Williamson (2003) have raised concerns in this regard. Boghossian sees these concerns as mainly stemming from an association of internalism with non-blindness – that is, with "reflectively appreciable warrants." He argues that all that the counterexamples to reliabilism show is an incompatibility between being blameless and being justified, and that this in and of itself does not support the claim that internalism requires accessible warrants. Furthermore, a rule-circular justification is

¹⁸⁶ He has in mind here the last part of *PI* 219:

When I obey a rule, I do not choose.
I obey the rule *blindly*.

Boghossian says that: "One of Wittgenstein's fundamental insights, it seems to me, was to realize that we must be capable of a form of blameless reasoning that did not depend on any other cognitive state of the thinker's and, in particular, not on anything analogous to sight." (2003: 277, fn. 18) In chapter 3, sections 3.3.2 and 3.3.3 we gave a different reading of this passage, where Wittgenstein's remarks on blindness are understood as making a more radical point: not as claiming that the thinker does not need to know anything about the epistemic status of the rule, but as claiming that the rule itself is irrelevant; it does not and it cannot provide guidance. Following the MPP rule is what we *say* we do, not what we actually do when it comes down to performing the basic inferential act itself.

our best (or maybe only) hope for justifying logical beliefs, but for it to work we need to give an adequate account of warrant-transmission that neither internalism nor externalism has been able to offer. So –

and here we see the elimination strategy at play again – there has to be a way of making sense of blind yet blameless inferential transitions if we are to avoid the ‘no justification is possible’ line of thought.

Let’s suppose an inferential transition that is both blind and blameless is what we need, both to avoid being irresponsible and to secure the transmission of warrant. How can an inferential transition come to be “blind yet blameless”? The possibility of relying on a list that simply takes which inferences fall into this category as a primitive doesn’t sit well with Boghossian’s concern about having something more substantive to say. A more pragmatic account, such as Harman’s ‘general conservatism’ – where beliefs and belief forming methods are *prima facie* justified until incoherence surfaces – is found dependent on a “conception of logical consequence and logical consistency,” and thus already presupposing an answer to the question. (See Boghossian 2003: 278, and Harman 1986: 29-42 and 2003: 25-26).

It is at this point that conceptual role semantics finally makes an entrance to save the day.¹⁸⁷ Boghossian writes:

A deductive pattern of inference P may be blamelessly employed, without any reflective appreciation of its epistemic status, just in case inferring according to P is a precondition for having one of the concepts ingredient in it. (2003: 279)

¹⁸⁷ Those who think that its introduction has taken too long would do well to remember Orson Welles’ remark about long-delayed entrances of main characters that, simply by such a move, become ‘star parts.’ He was of course alluding to his character Harry Lime in *The Third Man*. Marlon Brando’s Colonel Kurtz in *Apocalypse Now* would be another example – and we shouldn’t forget, as Michael Levin pointed out to me, King Kong and Keyser Soza.

And, bringing this home to our current problem, the point is that:

[...] we would have an explanation for the blameless blindness of MPP if it's constitutive of having the concept *conditional* that one *take* [my emphasis] p and ' $p \rightarrow q$ ' as a reason for believing q . (2003: 279)

So “blind yet blameless” inferences transfer warrants from the premise(s) to the conclusion, on the condition that the inferential transition itself is required for possessing one of the concepts involved in it. In Boghossian's version of the “traditional connection,” the link between concept possession and the epistemological conception of blind and blameless inference is such that: if it is a precondition of having the concept of the conditional that we take a belief in the premises ' $p \rightarrow q$ ' and ' p ' as a reason to believe the conclusion ' q ,' then we cannot be required to be in position to know *in advance* that the transition is truth-preserving, as that would *already* require having the concept of the conditional. Since that is exactly what we are claiming that we can't have without first taking the premises as a reason for the conclusion, the inference has to be blind. What is more, if the inferential move is constitutive of having the concept of the conditional, our taking the premises as a reason for the conclusion cannot be subject to accusations of epistemic irresponsibility (qualifications to follow), as no thoughts involving the concept of the conditional would otherwise be possible.

One of the questions that come to mind here is this: how is 'take' supposed to be understood when we say that the thinker takes ' p ' and ' $p \rightarrow q$ ' as a reason for believing ' q '? Boghossian's answer seems to be: “dispositionally.” Here are some relevant passages: “we come to grasp the logical constants by being disposed to engage in some inferences involving them and not in others” (2003: 279); “if fundamental inferential dispositions fix what we mean by our words, then

[...] we are entitled to act on those dispositions prior to and independently of having supplied an explicit justification for them” (2001a: 258); and “if an inferential disposition is meaning-constituting, then it is a fortiori reasonable, reasonably used independently of any belief about its properties.” (2001a: 265).¹⁸⁸

So, these passages suggest a dispositional reading. We take the premise(s) as a reason for the conclusion because we are disposed to do so. But things are a bit more complex, as we will see below.

Leaving dispositions aside for a moment, we should now consider Boghossian’s version of the principle linking semantics and epistemology:

(Meaning-Entitlement Connection or MEC): Any inferential transitions built into the possession conditions for a concept are *eo ipso* entitling. (2003: 241)

As stated, MEC seems quite straightforward. One of the first tests Boghossian has in store for it is overcoming the ‘bad company’ problem. And the prognosis is rather bad: as stated, MEC not only entitles us to the MPP rule but also to Prior’s rules for *tonk*, since the *tonk* ‘concept’ satisfies MEC just as much as the conditional. What is worse, *tonk* is not alone among the bad ‘concepts’ that MEC lets through without a problem. Dummett’s pejorative *boche* provides another example (Dummett 1973: 454):

¹⁸⁸ Also: “if it is really true that someone’s being disposed to reason according to modus ponens is a necessary condition of their having any logical concepts at all, and so of being able to reason in any shape manner or form, there can be no intuitive sense in which their disposition to reason according to modus ponens can be held to be irresponsible, even in the absence of a reflectively appreciable warrant that justifies it.” (2001a: 260)

x is German

x is *boche*

x is *boche*

x is cruel

It seems that being straightforward is not helping MEC all that much. A restriction is certainly needed to filter out cases such as *tonk* and *boche*. Two possibilities come under Boghossian's consideration: restricting the notion of a genuine concept to concepts that are constituted by truth-preserving rules, or modifying the rules themselves by making them conditional.

One of the problems with the first alternative is that, though it filters out both *tonk* and *boche*, it lacks plausibility. While a case like *tonk* might suggest that there is no genuine concept involved, with *boche* the intuitive diagnosis is not that the concept is defective, but that its rules don't provide us with entitlement.¹⁸⁹ After all, pejoratives like *boche* are used, and having to deny them a genuine concept status to block warrants for their use doesn't seem an acceptable reconstruction of what is going on with such cases.

An additional trouble for this alternative is that a concept we would not count as genuine will pass as genuine, provided its rules are truth-preserving. One of the examples offered by Boghossian is the concept *flurg*. Suppose someone introduces it by stipulating the following introduction and elimination rules:

¹⁸⁹ Horwich (2005: 153-154) claims, however, that neither the rules for *tonk* nor the rules for *boche* are genuine concept-constituting rules, and so they don't constitute genuine concepts. The rules for *tonk* cannot be completely followed because, since *tonk* would lead us to believe every proposition, including any instance of 'p and not p,' they are inconsistent with the concept-constituting rules for 'not.' Similarly, the rules for *boche* are inconsistent with the rules for cruel (Horwich uses 'English,' 'Pom,' and 'pretentious' in his version of Dummett's example, but his point would also apply to the original *boche* example). Peacocke also denies that *tonk* constitutes a genuine concept (1992: 21).

x is an elliptical equation

x is *flurg*

x is *flurg*

x can be correlated with a modular form

The Taniyama-Shimura conjecture, that every elliptical equation can be correlated with a modular form, was proved in 2001 and is currently known as the modularity theorem. The rules for *flurg* are thus necessarily truth-preserving. Yet, there is no doubt that the mere introduction of a concept by stipulation can give no entitlement to the conjecture (which was in fact quite hard to prove, Wiles first proved it for semistable elliptical curves, the case needed to prove Fermat's last theorem, and later on it was possible to prove it in full generality). What we have here is therefore a case with necessarily truth-preserving rules and either no genuine concept or a defective concept resulting from them.

These problems lead Boghossian to a "change of heart" (2003: 243 fn. 24) regarding truth-preservation, a requirement he defends in (2000 and 2001a).¹⁹⁰ The second option, proposed in his 2003 paper, is quite a bit less straightforward than the original MEC. It depends on introducing 'conditionalized' possession conditions for some concepts, and on a distinction between defective and non-defective concepts.

¹⁹⁰ Boghossian was following Peacocke (1992 and 1993) in requiring the constituting rules to be truth-preserving for a concept to be a genuine concept. Note that in Peacocke's account, pure semantical considerations guarantee that the constitutive rules of a genuine concept cannot fail to be truth-preserving. As Williamson (2003) observes, this makes the proposal dangerously close to externalism.

Later on, Peacocke modifies his view and replaces the original concept possession conditions by "implicit conceptions," which are meant to be more flexible and, among other things, waive any explicit endorsement of the concept-constituting rules and propositions (see 1998a, 1998b, 1999, and 2004). More recently, in another and more drastic modification, Peacocke claims that a concept is not individuated by its conceptual role but by its 'fundamental rule of reference.' Briefly, the fundamental rules specify the reference conditions of the concepts in such a way that the rationality of believing a proposition is explained by the rationality of believing the likelihood that the proposition's truth-conditions will be satisfied. (See Peacocke 2008)

The thought behind the conditional version is that concept possession cannot depend on having particular views or submitting to a certain theory. We cannot simply posit factual claims as being constitutive of a concept and hold that only those having the relevant beliefs possess the concept. The possibility that there isn't anything falling under the concept must remain open, no matter how certain we might be concerning the beliefs in question. We might be wrong and, in any case, concept possession should leave space for disagreements about the extension of the concept.

Consider *flurg*. In view of the given rules, to possess the concept we would have to believe that for every elliptical equation there is a correlative modular form. Such stipulation violates the intuitive constraints just discussed. Instead, a conditional version of *flurg* allows for the possibility that there might be no property that satisfies the rules for the original version of *flurg*. Boghossian's conditional version of the stipulation is:

If there is a property which is such that, any elliptical equation has it, and if something has it, then it can be correlated with a modular form, then if x has that property, x is *flurg*. (2003: 284)

He does not deny concept status to the unconditional version of *flurg* – concepts are, he says, relatively cheap – but claims that it makes for a defective concept. The conditional version of *flurg* lets us both have the concept and deny that every elliptical curve can be correlated with a modular form because it contemplates the possibility that there is no property that can play *flurg*'s role. And for this reason, conditional *flurg* can be considered a non-defective concept.

According to Boghossian, not all concepts have conditional versions. In particular, MPP does not have a conditional version. Since introducing a concept via a conditional stipulation already

requires having the concept of the conditional, only an unconditional version is available for the concept of the conditional. The same holds, in his view, for other “primitive logical constants,” and presumably for the quantifiers, since conditional versions presuppose the ability to “refer to and quantify over properties in particular, and semantic values, more generally.” (285) In general, Boghossian holds that whenever a conditional version is possible, the unconditional version results in defective concepts, while the conditional version always results in non-defective concepts. Finally, if only an unconditional version is possible, the resulting concept is non-defective, given that:

In such a case, there is no alternative but to accept ‘conditional theory’—Modus Ponens in effect—if you are so much as to have the conditional concept. It thus couldn’t be epistemically irresponsible of you just to go ahead and infer according to Modus Ponens without conditionalizing on the existence of an appropriate truth function for it—that is simply not a coherent option in this case. (2003: 285)

With the proper fixes in place, the new version of MEC reads as follows:

Any rules that are written into the possession conditions for a *non-defective* concept are *a fortiori* entitling. (2003: 285)

Boghossian’s answer to ‘bad company’ then goes as follows. The problem with *boche* and *flurg* is that they were introduced by unconditionalized versions, when conditional versions would have been possible. They are, therefore, defective concepts. *Tonk*, on the other hand, doesn’t constitute a genuine concept because its rules cannot be followed thoroughly. Finally, MPP inferences are warrant-transferring because the MPP inference itself constitutes the non-defective concept of the conditional.

We can now go back to the justification of the MPP belief and use a rule circular argument to derive the associated conditional of the MPP rule. Such argument will make one inferential step according to the rule, but that should not be reason for concern because MPP inferences have been shown to be warrant-transferring and, hence, our entitlement to them has been secured.

What about ‘begging the question’? A rule-circular argument remains powerless to persuade the skeptic. The reason is that it will be open to the skeptic to relocate his doubts on the validity of MPP to the concept-constituting character of the introduction and elimination rules for the conditional, which in turn will call into doubt the transfer of warrants in MPP inferences.

Yet, Boghossian thinks he has fulfilled the demand of showing why it is okay to use MPP in an inferential step without having earned a justification for its validity. Moreover, he points out that what a rule-circular argument ultimately achieves is anything but insignificant; it secures a justification of the belief that all instances of MPP are truth-preserving. We won’t be able to persuade a skeptic, but nevertheless we have acquired a justification.

6.4 Boghossian’s Proposal and the Adoption Problem

Boghossian’s defense of rule-circular arguments and blind yet blameless inference has been subject to various criticisms which target different aspects of the proposal. Among the most determined critics of his account, we find Horwich (1997, 2000, 2005) and Williamson (2003, 2007: ch.4), though Wright (2001 and 2004b) also raises a number of objections.¹⁹¹ Without a doubt, the problems of ‘bad company’ and ‘begging the question’ (both acknowledged by Boghossian himself from the outset), give critics a fertile ground to work in. But Boghossian sees

¹⁹¹ See Besson (2009), Schechter and Enoch (2006), Fodor and Lepore (1991 and 2006).

his “neo-rationalist program” as a work in progress, and despite the critics’ efforts, the objections seem to have left his confidence on the prospects of the proposal basically unmoved.

Williamson has tirelessly pointed out (see 2003, 2007, 2011, and 2012) that the link between accepting a pattern of inference and possessing a concept fails and he has given several counterexamples of it. Moreover, he argues that the rules for *boche* are not concept constituting, since we can possess the concept without following its introduction and elimination rules. As these are pejorative concepts, we may possess them but, for obvious reasons, be unwilling to use them.¹⁹² Williamson expresses doubts that we can come to possess new concepts by introducing stipulations such as the one for *flurg*.

Both Williamson and Horwich argue that it is possible to have a conditional version of the logical constants (Horwich 2005: 150; Williamson 2003:287-288). Horwich holds that Boghossian cannot explain the obligation to accept concept-constituting beliefs, but just our right to accept them (Horwich 2000: 161; 2005: 157, 163). He has also argued that no account of our entitlement for employing a basic rule such as MPP is needed, since it is explanatorily fundamental that we are justified in our use of MPP and other rules (2005: 169-170).

Wright’s main concerns revolve around the problem of begging the question and the objectivity of logic. In Wright’s view, if what we are interested in is not just a justification of our knowledge of the validity of MPP but of our *claim* to know this,¹⁹³ Boghossian’s account leaves us without an answer. According to Wright, for Boghossian:

¹⁹² Williamson’s own account of pejoratives relies on the notion of conventional implicature to explain their derogatory status which, for him, cannot be properly explained in the conceptual-role model. See Williamson 2003 and 2006.

¹⁹³ Wright has stressed this distinction and argued that externalist positions are unable to account for our claim to know (Wright: 2004a).

a derivation – a rule-circular derivation – of a schematic statement of MPP, say, may in principle be at the service of the acquisition of knowledge that the rule is valid even though that very rule is used therein. If the question is not how we might get such knowledge but our right to claim it, then Boghossian’s proposal is no help – that the claim to knowledge must rest, even in the rule-circular case, on a prior claim that the machinery utilized in the rule-circular derivation is valid. (Wright: 2004b)

Wright connects this point with the problem of the objectivity of logic. His worry is that if a rule-circular argument cannot make good on our claims to knowledge (and thus move the skeptic), the threat of logical relativism will remain looming.

Some of these objections are difficult to deal with (or, to be frank, to even see how they could be dealt with). Still, in his replies to some of these criticisms we find little sign of substantial doubt in Boghossian’s commitment to the “neo-rationalist program.”¹⁹⁴ The situation changes, however, in his more recent paper “What is Inference?” (Boghossian 2014) (already foreshadowed by Boghossian 2008a), where we find him giving room to a rather pessimistic attitude towards his own program, and perhaps even to the prospect of giving a justification for logical beliefs at all. The objection motivating the now pessimistic attitude is put forward by none other than Boghossian himself. There are important points of connection between some of the problems Boghossian (2008 and 2012a) raises and the issues related to the adoption problem. In what comes next, we will explore this connection and how it bears on Boghossian’s proposal.

What can one say, then, about Harry from the point of view of Boghossian’s proposal? One could expect that the concept of ‘all’ would be constituted in a similar way to that of *boche*, *flurg*, and the propositional constants, that is, by means of its introduction and elimination rules:

V-Elimination or Universal Instantiation

¹⁹⁴ See, for example, Boghossian (2008a, 2011, 2014)

$$\underline{Vx Px}$$
$$Pa$$

V-Introduction or Universal Generalization:

$$\underline{Pa}$$
$$Vx Px$$

where x does not occur free in any premise which has been used in the proof of Pa .

Now, we said that on this account, for a thinker to possess the concept, she has to *take* the premises of these inferences as a reason for believing the conclusion. We also saw that, for Boghossian, dispositions play an important role. He claims that we grasp the logical constants by being disposed to engage in some inferences (2003: 279); that (some) inferential dispositions fix the meanings of the logical constants (2001a: 258); that if an inferential disposition is meaning-constituting, it is justified without argument, etc. Sometimes, he expresses the point in terms of rule-following: “following certain inferential rules is constitutive of our grasp of the primitive logical constants” (2003: 286), but rule-following is understood in dispositional terms, “as a disposition to rule-conform under appropriately idealized circumstances.” (2000: 230)¹⁹⁵ So, for example, for someone to be disposed to infer according to MPP, she has to be disposed, whenever she believes that p and that ‘if p then q ,’ to believe q . (See 2000: 229, and 2001: 633)

Now, if it is a precondition of having the concept ‘all’ that we take a belief in the premise $Vx Px$ as a reason to believe the conclusion Pa , and the content of ‘all’ is fixed by means of inferential

¹⁹⁵ “By saying that a thinker has and operates according to these rules I don’t mean that the thinker grasps these rules as propositions. I mean that he follows these rules, and that this shows up in his behavior, however exactly that is to be analyzed. It will do no harm, for present purposes, to think of rule-following as a disposition to rule-conform under appropriately idealized circumstances.” (2001a)

dispositions, then it would seem that it is a precondition of having the concept ‘all’ that we are disposed, whenever we believe $\forall x Px$, to believe Pa .

For Harry, this doesn’t look very promising since he obviously lacks the relevant disposition. Recall the intuition we had at the beginning of this chapter: if, after we explicitly tell him the relevant rules, he is still unable to infer accordingly, then he does not understand what ‘all’ means. Boghossian could easily explain this: we can say that Harry does not possess the concept ‘all,’ given that to possess it he would have to be disposed to follow the UI rule, that is, disposed to believe Pa from $\forall x Px$. Since this is not the case, he does not possess the concept ‘all,’ and it is thus hardly surprising that he does not understand what it means or doesn’t make the inference.

Is this diagnosis correct? On a first pass, it seems to give a reasonable explanation of what is going on. But to evaluate it properly we need to have a clear picture of how acceptance, understanding, and the disposition to follow the rule come together. There seem to be three main possibilities:

- (a) The disposition to follow the rule comes first, while the thinker’s acceptance of the rule and understanding of ‘all’ is implicit in her exercise of the disposition.
- (b) Acceptance of the rule, understood as an explicit or implicit belief that the rule is correct, comes first. The thinker’s disposition to follow the rule and understanding of ‘all’ are a consequence of such acceptance.
- (c) Acceptance of the rule (again, understood as an explicit or implicit belief) and understanding come first, and together. The disposition to follow the rule is grounded on such acceptance/understanding.

The explanation we gave of Harry's situation coincides with (a).

If acceptance is not understood as either an explicit endorsement of the statement of the rule or an implicit mental state encoding the content of the rule, but is simply exercised as an actualization of an inferential disposition, then it is clear that there is nothing that we can *say* to Harry that would help him to acquire the concept 'all.' But is this what Boghossian has in mind?

There are reasons to believe that this is not Boghossian's view, and that this is incompatible with the role that implicit definitions play for concept-based accounts and the epistemological upshot that such accounts seek to obtain.

Let's first consider why this is not Boghossian's view. To see why it is not merely a disposition to follow the rules but the acceptance of the rules themselves what is doing the relevant work, it is instructive to consider the introduction of concepts like *flurg* or *neutrino*. For his treatment of such concepts, it is clear that the rules are meant to state what someone needs to believe to acquire the concepts. And he explicitly says so in his reply to Williamson (2003):

what I want to capture is the idea of someone who is only prepared to infer according to the boche rules because they antecedently believe that 'There is a property F , such that $Gx \rightarrow Fx$ and $Fx \rightarrow Cx$ '. And I don't see how their reasoning could depend on that without their having, at a minimum, the conceptual materials that make up the antecedent of the Carnap sentence, including the quantificational apparatus and the conditionals that make up the statement of the theory. If all of this is right, it follows that conditional counterparts for one's primitive logical constants will not be available and hence that one could hardly be blamed for employing their unconditionalized versions. In particular, you couldn't conditionalize on the existence of an appropriate truth function for the conditional, for you would need it in order to conditionalize on anything. In such a case, there is no alternative but to accept "conditional theory" – modus ponens and Conditional Proof, in effect – if you are to so much as have the conditional

concept. It thus couldn't be epistemically irresponsible of you to just go ahead and infer according to MPP without conditionalizing on the existence of an appropriate truth function for it – that is simply not a coherent option in this case. (Boghossian 2003: 287)

There is then a “theory” that needs to be accepted for someone to have the concept of the conditional. Although in most cases acceptance of the rule would be merely implicit in the thinker's practice of inferring according to the UI introduction and elimination rules, if the rules are to be said to constitute the concept, then it must be possible to come to grasp and possess the concept by accepting them. So someone who does not possess the concept should be able to acquire it on the basis of the rules.

From this it is clear that if (b) is the option that correctly characterizes Boghossian's account, one of the reasons there cannot be dispositions that are not grounded on acceptance are not a viable option for justifying concept-constituting is given by Boghossian himself:

If all I can say about my disposition to reason according to MPP is that that is how I am disposed to reason, what is to rationally prevent someone else from proceeding in some other way? (2001c: 634)

Ungrounded dispositions would relegate the role of the rules to mere descriptions. In concept-based accounts, however, rules are supposed to have a constitutive role. And because of that, because they represent the conditions for possessing the concept, we are entitled to use them. On the view that they simply describe our inferential dispositional practice, two problems that concept-based accounts were meant to address come right back to haunt us. One is the problem that Boghossian mentions above, for rules would be merely descriptive of the practice and, hence, powerless to impose normative descriptions. The other is that concept-based accounts

were meant to provide a means of selecting the dispositions that would constitute the concept and ground them on prior acceptance of the rules, securing at the same time our entitlement to them. But if the rules describe inferential dispositions, how is our entitlement to them going to be accounted for?

Now, if acceptance of the rules is going to come first, and if it is in virtue of them that we come to possess the concept and be able to exercise our inferential dispositions, where does this leave Harry? It doesn't take much to see what the problem is going to be.

Even though in most cases thinkers would not have any explicit acceptance of the rules, as they would be implicit in their inferential practices and probably picked up by being participants of that practice, for the concept-constituting theory to be adequate it should be in principle possible to acquire the concept by accepting the rules. If this were not possible, it would suggest that it is the use of the rules in practice that is constituting the concept and not the implicit definitions given by the rules. Could we get Harry to acquire the concept by accepting the rules? We have tried this before. It would be in principle impossible for him to follow the rules while lacking a prior ability to perform universal instantiations.

But if this is the model, the problem is not simply that it presupposes what it is meant to justify. The deeper problem is that this model makes the very possibility of acquiring the ability to perform inferences according to basic rules impossible. And the moral of our discussion seem to go even further, because it appears to follow from it that Boghossian's principle LP, and his taking condition, will have to be dropped.

Chapter 7

On the Nature of Basic Inferences

The main moral of the adoption problem appears to be a negative one: basic rules of inference do not play a fundamental role in shaping actual inferential uses.¹⁹⁶ What we have found problematic is not a particular characterization of *how* they are accessed or justified, but the very assumption that it is *because* we have accepted, internalized, or somehow implicitly accessed their content that we can take a conclusion to follow from the premise(s) in an inference. We cannot think of them as guiding or instructing (either explicitly or implicitly) our inferring. If they did guide, Harry could be helped by our explicitly providing them. Yet, only if we already infer may they be of use.

If it is *impossible* for someone to infer from mere consideration of basic logical principles or rules, the capacity or ability to infer has to be conceptually prior. Without it, propositional knowledge would be inert: neither principles nor rules are inference tickets for getting to the conclusion. But then, it is not *because* we somehow rely on rules of inference that we infer according to them; quite the opposite, it is because our deductive inferences can be said or thought to be in accordance with them that we can recognize them and we may find their formulation useful.

¹⁹⁶ I'm ignoring here the possibility that rules may be hardwired, which will be discussed in section 7.2.1

Versions of what are perhaps the three historically more influential views on the justification problem¹⁹⁷ – based on the web of belief model, on rational intuition, and on concept-constitution – have been discussed in the previous chapters. Our main argument against them has been that precisely because our acceptance or internalization of basic principles and rules is given or assumed to have a fundamental role in these proposals, the nature of the inferential process is distorted to the point that it is hard to see how inferring could be possible at all. Moreover, if the capacity or ability to infer is conceptually prior and condition of possibility of the formulation and acceptance of the logical rules and principles our inferential acts are said to conform to, by misconceiving the nature of basic inferences these proposals undermine as well the rules and principles they seek to justify.

We should emphasize that the problem here is general, it does not depend on the specific justification proposals we have discussed: the point is not that there will be inferential transitions that are not covered by the justification proposal on offer. It really doesn't matter what kind of justification we favor, whether we claim that logical principles are knowable *a priori* or *a posteriori*, or whether we claim that our knowledge of them is unjustifiable. Even their truth or validity is beside the point. The issue is what role these rules can have in actual instances of inferring.

It is for this reason that it was said in the beginning that we need to have a better understanding of the nature of inferring before we attempt to justify the belief in logical principles or rules, or try to give an account of our entitlement to use those inferences. On pain of risking an implausible account of what inferring according to basic principles is, we cannot undertake

¹⁹⁷ Recall that in chapter 1 we followed Wright's characterization of the three main problems of the epistemology of logic: the justification problem (about logical beliefs), the entitlement problem (about the use of inferences that accord with rules of inference), and the problem of the nature of inference.

the justification. And only by getting clearer on this prior problem we may be able to get a better grip on the conceptually separable problem of justification.

One possibility that opens up once we accept that propositional content doesn't play a fundamental role in the basic inferential processes themselves is to understand inference as a species of *knowing how*. The idea is far from new. Indeed, Ryle's famous defense of the distinction between *knowing how* and *knowing that* specifically challenges the idea that inferential acts can be analyzed in terms of propositional knowledge. What makes Ryle's view particularly interesting for us is that it is based on his interpretation of Carroll's note as claiming that performing basic inferences cannot be a matter of *knowing that*. In the next section, we will discuss Ryle's arguments for the case of logic and explore the possibility that inferring is a kind of knowing how.

7.1 Ryle on Knowing How/ Knowing That

Ryle's paper "Knowledge How and Knowledge That" (1945) and his chapter 2 of *The Concept of Mind* (which seems to be a reworked version of this paper) introduce and defend the distinction between knowledge how and knowledge that. The main point Ryle wants to establish is that knowledge how cannot be understood as a species of knowledge that. Those who hold the opposing view are subject, he says, to the "intellectualist legend," a view that in his opinion hardly makes sense.

Ryle takes the distinction between knowing that something is the case and knowing how to do something to be a familiar one. And indeed, the distinction has been tacitly assumed in diverse areas and has gone mostly unchallenged until fairly recently, when Stanley and Williamson (2001) proposed an original account of knowledge how in terms of propositional knowledge and

criticized some of Ryle's arguments against the so-called intellectualists (those that, in his view, accept the intellectualist legend). As the fast growing list of contributions to the debate indicates, the debate between intellectualists and anti-intellectualists has remained very much alive ever since. The contrast between attitudes towards the distinction after and before Stanley and Williamson (2001) is quite striking: nowadays, everything about it appears to be subject to heated disputes, including the interpretation of Ryle himself. But the disputed issues transcend, of course, the interpretation of Ryle, and the focus has mainly settled on the question whether Stanley and Williamson's intellectualist account of knowing how can do the work it is intended to do, and whether an anti-intellectualist construal of knowing how and skillful action should be given in relation to dispositions, abilities, or something else (and how to delineate these notions).¹⁹⁸

What is (somewhat) more surprising is that Ryle's arguments in favor of the distinction, which had been (at least tacitly) accepted for a long time, appear now to be much less transparent than originally thought.¹⁹⁹ Indeed, alternative interpretations of both Ryle's arguments for the distinction and his characterization of knowledge how are currently quite common.²⁰⁰ Having other fish to fry here, we'll try to avoid trouble and keep to the basic ideas required to introduce Ryle's interpretation of Carroll. Because the discussion of Carroll is omitted in *The Concept of*

¹⁹⁸ There are of course many more issues going on in the debate. For an overview, see the introduction to Bengson and Moffett (2012).

¹⁹⁹ Leaving aside the change of attitude, Ryle's witty and almost outraged tone, as if he couldn't believe that this intellectualist legend has been allowed to stand for so long, has probably not helped. Although Ryle's prose is quite straightforward, his vehement denunciation sometimes seems to get in the way of giving detailed arguments and explaining the comparisons he is relying on. We will see below a case of the latter.

²⁰⁰ Hornsby (2012: 82), for example, denies that Ryle offered a "definite account" of know how, and Weatherson (2007) claims that Ryle's account of knowing how cannot be identified with dispositions to ϕ . See Hawley (2003) for an account of knowing how as ability under normal counterfactual conditions. See also Stanley (2011: chapter 1), Snowdon (2003, 2012), and the introduction to Bengson and Moffett (2012).

Mind, we will mostly concentrate on the 1945 paper. As we will see, a lot of what Ryle has to say – and this is a feature that perhaps hasn’t been sufficiently emphasized – has a distinctive Wittgensteinian flavor, in particular when it comes to his views on rules. Let’s start with Ryle’s characterization of the intellectualist position and his much discussed regress argument.

Ryle’s complaint against the intellectualists revolves around the role they assign to propositional content. According to him, what in this model distinguishes intelligent “practical activities” from non-intelligent ones is that the former, but not the latter, presuppose “contemplating” or “considering” propositions. (1945: 1-2) Propositional knowledge (knowledge that) is assumed to have a guiding role with respect to our intelligent practical actions and skills, which in turn are only intelligent in a derivative way. A prior grasp of propositional content is hence needed for performing intelligent acts, and knowing how is assimilated to a kind of knowing that.

Ryle thinks that, quite to the contrary, no such acknowledgment of propositional knowledge is required for intelligent action or skill, and that both practical and theoretical acts can be directly intelligent: “whatever ‘applying’ may be, it is a proper exercise of intelligence and it is not a process of considering propositions.” (1945: 3) In his view, not only is knowledge how irreducible to knowledge that, but it is “a concept logically prior to the concept of knowledge-that.” (1945: 4-5)

Ryle claims that his case against the intellectualists basically rests on a variation of one argument.²⁰¹ He says:

²⁰¹ As many authors have pointed out, it is difficult to see some of the other arguments Ryle gives as a variation of the regress argument. Bengson and Moffett (2012: 11ff) identify three of them: (1) that intellectualism doesn’t do justice to the intuitive and familiar distinction between theoretical and practical knowledge (1945: 4); (2) that ascriptions of intelligence or stupidity are problematic on an intellectualist view because the relation between mental states and particular performances is not accessible to an

The prevailing doctrine leads to vicious regress, and these in two directions. (1) If the intelligence exhibited in any act, practical or theoretical, is to be credited to the occurrence of some ulterior act of intelligently considering regulative propositions, no intelligent act, practical or theoretical, could ever begin. [...] (2) If a deed, to be intelligent, has to be guided by the consideration of a regulative proposition, the gap between that consideration and the practical application of the regulation has to be bridged by some go-between process which cannot by the pre-supposed definition itself be an exercise of intelligence and cannot, by definition, be the resultant deed. This go-between application-process has somehow to marry observance of a contemplated maxim with the enforcement of behaviour. [...] Consistency requires, therefore, that this schizophrenic broker must again be subdivided into one bit which contemplates but does not execute, one which executes but does not contemplate and a third which reconciles these irreconcilables. And so on for ever. (1945: 2-3. See also 1949: 30)²⁰²

The basic structure of the argument seems to be the following. Suppose that for an act to be intelligent it has to be guided by propositional content. Prior grasp or contemplation of the proposition would then be needed to initiate an intelligent action. But applying propositional knowledge to an action is itself an action, and for that action in turn to be intelligent it would also have to be guided by propositional content, thus a prior grasp of such content would also be required, and so on *ad infinitum*.²⁰³

For Ryle, the application of a rule or maxim is not straightforward, it is something that can be done “intelligently” or “poorly.” Consider the case, which will be important for us later, of the chess players:

external observer (1949: 21); and, finally, (3) that a fool might have all the relevant propositional knowledge without knowing how to perform. (1945: 8)

²⁰² In *The Concept of Mind*, Ryle introduces the regress argument in the following way: “The consideration of propositions is itself an operation the execution of which can be more or less intelligent, less or more stupid. But if, for any operation to be intelligently executed, a prior theoretical operation had first to be performed and performed intelligently, it would be a logical impossibility for anyone ever to break into the circle.” (1949: 30).

²⁰³ See Bengtson and Moffett (2012: 8ff) for discussion of this argument.

[...] what truths does the clever chess-player know which would be news to his stupid opponent? Obviously there is no truth or set of truths of which we could say, 'If only the stupid player had been informed of them, he would be a clever player,' or 'When once he had been apprised of these truths he would play well.' We can imagine a clever player generously imparting to his stupid opponent so many rules, tactical maxims, "wrinkles," etc., that he could think of no more to tell him; his opponent might accept and memorise all of them, and be able and ready to recite them correctly on demand. Yet he might still play chess stupidly, that is, be unable intelligently to apply the maxims, etc. (1945: 5)

So propositional knowledge is insufficient for an intelligent performance. What is missing is knowledge when and how such propositional knowledge should be applied for the performance to be successful.

Ryle's regress argument has been called into doubt by Stanley and Williamson (2001; see also Stanley 2011a and 2011b: ch. 1). They reject Ryle's claim that the intellectualist is committed to maintaining that an intelligent action has to be preceded by the act of contemplating or considering the proposition meant to guide the action. According to Stanley, "it is just *manifestly absurd* that all intelligent actions are preceded by distinct actions of self-avowals of propositions" (2011a: 14). In their view, it is possible to hold that an action is intelligent in virtue of being guided by a rule or maxim and, at the same time, deny that such guidance is the result of an act of contemplating the proposition.²⁰⁴ A mental state encoding propositional knowledge can be exercised without the need of conscious consideration of its content. Moreover, he argues, Ryle simply assumes, without offering a supporting argument, that it is possible to exercise knowledge

²⁰⁴ They quote a passage from Ginet where it is argued that Ryle's regress only shows that a prior internal act of considering a proposition and then inferring the relevant instructions is not required for action. He says: "I exercise (or manifest) my knowledge *that* one can get the door open by turning the knob and pushing it (as well as my knowledge *that* there is a door there) by performing that operation quite automatically as I leave the room; and I may do this, of course, without formulating (in my mind or out loud) that proposition or any other relevant proposition" (Ginet 1975: 6-7).

how directly, without relying on any mental state of triggering a representation of the relevant knowledge. But, in his view, there is no reason to hold that there's an asymmetry between knowledge how and knowledge that on this issue. If knowledge that requires a mental act of triggering a relevant representation to be manifested, so does knowledge how. In both cases it is required, but in neither of them does the content need to be consciously accessed by the agent.

If manifesting one's propositional knowledge that one can open the door by turning the doorknob requires a prior triggering of a representation of the doorknob, why doesn't manifesting one's knowledge of how to open the door require a prior triggering of a representation of the doorknob? That is, if knowing that requires prior triggering of representations, why doesn't knowing how require prior triggering of representations? (2011a: 16)

In Stanley's opinion, what Ryle's regress argument shows is that "the knowledge that we act on when we act intelligently sometimes is manifested *directly*, without a prior mental action" (2011a: 19). We will come back to some of the points raised here. But let's consider first Ryle's view according to which performing deductive inferences does not require propositional knowledge, but knowing how to infer.

Ryle draws on his reading of Carroll's note to show that regress ensues if propositional knowledge is assumed to have a guiding role in the inferential process. This particular regress, let's call it 'the logical regress,' is supposed to be a special case of the general regress argument. As was the case with Kripke, Ryle replaces Carroll's example and this modification affects his reading of the note. He writes:

A pupil fails to follow an argument. He understands the premises and he understands the conclusion. But he fails to see that the conclusion follows from the premisses. The teacher thinks him rather dull but tries to help. So he tells him that

there is an ulterior proposition which he has not considered, namely, that if these premises are true, the conclusion is true. The pupil understands this and dutifully recites it alongside the premisses, and still fails to see that the conclusion follows from the premises even when accompanied by the assertion that these premisses entail this conclusion. So a second hypothetical proposition is added to his store; namely, that the conclusion is true if the premises are true as well as the first hypothetical proposition that if the premisses are true the conclusion is true. And still the pupil fails to see. And so on forever. He accepts rules in theory but this does not force him to apply them in practice. He considers reasons, but he fails to reason. (This is Lewis Carroll's puzzle in "What the Tortoise said to Achilles." I have met no successful attempt to solve it.) (1945: 6)

And the moral he draws is the following:

What has gone wrong? Just this, that knowing how to reason was assumed to be analysable into the knowledge or supposal of some propositions, namely, (1) the special premisses, (2) the conclusion, plus (3) some extra propositions about the implication of the conclusion by the premisses, etc., etc., ad infinitum. "Well but surely the intelligent reasoner is knowing rules of inference whenever he reasons intelligently." Yes, of course he is, but knowing such a rule is not a case of knowing an extra fact or truth; it is knowing how to move from acknowledging some facts to acknowledging others. Knowing a rule of inference is not possessing a bit of extra information but being able to perform an intelligent operation. Knowing a rule is knowing how. It is realised in performances which conform to the rule, not in theoretical citations of it. (1945: 6-7)

Here Ryle is taking issue with a case of intelligent action *par excellence*, deductive inferential acts, which was "supposed to be one of the opposition's strongholds" (Rumfitt 2012: 335). Yet, on Ryle's view, performing a deductive inference is not a matter of knowing that a proposition about the implication relation between premises and conclusions is true or knowing that a rule is valid, but rather of knowing how to 'move' from the premises to the conclusion.

Stanley (2011a), however, interprets Ryle's logical regress argument as being about what we have called 'the justification problem.' For him:

The purpose of Ryle's regress argument in *The Concept of Mind* is to refute the view that intelligent action is action under the direction of the right kinds of propositional states. The purpose of Ryle's version of the Lewis Carroll regress is to refute the view that *epistemic justification* is entirely a matter of having the right propositional states. The moral of both arguments is that propositional knowledge is not enough to explain intelligence, whether practical or theoretical. (Stanley 2011a: 27; my emphasis.)

The upshot is that “knowing how to apply a rule is not an instance of knowing that something is the case” and so “epistemic justification is not just a matter of knowing truths, but involves [...] knowing how to perform the inference,” as for Ryle, “we are justified in accepting an inference only if we know the premises, and we know how to perform the inference” (2011a: 28).

Why does Stanley think that Ryle's reading of Carroll emphasizes the justification problem? Ryle seems to be addressing a more mundane issue about inferential acts. In line with what he had been arguing before, his concern is with actions and particular performances, with whether knowing how to do something can be understood in terms of knowing that something is the case. In this particular case, the question is what is needed for the pupil to perform the inferential act, and not what justifies him in performing it. As Stanley himself notes, Ryle's pupil is not a skeptic that is seeking assurances every step of the way, he is not at all claiming that the inferential transition lacks justification, he just doesn't realize that the conclusion follows from the premises. Moreover, the moral Ryle draws from the example is explicitly about the inferential process itself, as he makes it clear by saying that “...knowing how to reason was assumed to be analysable into [...]”. Simply maintaining that knowing a rule of inference is “knowing how to move from acknowledging some facts to acknowledging others” doesn't say anything about whether we are justified in doing so.

On this point, Ryle's reading of Carroll is similar to Kripke's. Both set skeptical concerns aside and introduce the case of someone who does not get that the conclusion follows from the premise(s). This brings the center of attention to the inferential act itself: what would it take for such a person to draw the conclusion? But Harry's case seems to be more extreme, as he *never* inferred in accordance with UI. We are not told whether the pupil has drawn similar inferences in the past, not even what the inference in question is supposed to be, all we know is that he just can't see in *this* particular case that the conclusion follows.

There is, however, a difference in what we told Harry and what the teacher tells the pupil. We gave Harry a statement of the UI principle and of the UI rule of inference in the hope that they would help him perform an inference in accordance with UI. In contrast, the teacher merely tells the pupil "that if these premises are true, the conclusion is true." One might wonder how such statement could be of any help to the pupil, even if we assume that the inference is valid (Ryle doesn't say whether this is the case). The teacher's intervention would be more appropriate for the case of someone who *does see* that the conclusion follows but still *doubts* that the inferential transition is justified. (This is perhaps the reason why Stanley reads Ryle as raising a problem for the justification of the inference.)

However, the teacher's intervention is just inappropriate considering the pupil's problem. How could such a general statement help him 'see' anything? Wouldn't it be more appropriate to tell him the rule or rules of inference that he needs to apply to reach the conclusion? Of course, if he never inferred in accordance with them, that won't help either. But we are assuming that he has indeed performed similar inferences (Ryle doesn't say anything to the contrary).

Imagine that the inference in question is one of those proposed by Boghossian, such as the one about Fermat's last theorem or the Taniyama-Shimura conjecture. If we assume that the pupil

has to prove a case of the conjecture, that if x is an elliptical equation, then x can be correlated with a modular form, would we think that the pupil is 'dull' because he doesn't 'see' that it follows? If anyone is 'dull' in this scenario, it would surely be the teacher for thinking that merely saying that "if the premise is true, the conclusion is true" could possibly help. Ryle's example is under-described, but even assuming that we are dealing with a simple inference, the information the teacher provides is irrelevant to the pupil's problem. What could help him is an explanation of why it follows, not to be told that it does, and stating the relevant rule or rules of inference *might* well be a step in the right direction.

Likewise, if the point Ryle wants to make is that in the context of an inference someone may consider a rule but fail to apply it, giving an example of a specific rule of inference would have served his purpose better. For no rule asserting the conclusion to be true when the premises are true needs to be applied to derive a conclusion in the context of an inference. Ryle was, of course, following Carroll on this point, but once the example has been changed the appeal to the validity of the inference loses its relevance.

That Ryle thought the moral of the logical regress was about the application of a rule or maxim is confirmed by his treatment of the pupil's case in parallel to the case of the stupid chess player. Both the pupil and the stupid chess player fail to act on rules after considering their content, and Ryle's point is that grasping them does not tell us which rule to apply at the appropriate moment or how to follow it. "The application of maxims, etc., is certainly not any mere contemplation of them." (1945: 6). Applying a maxim or a rule is itself something that can be *done* poorly or intelligently. So, both cases are meant to establish that knowing how to apply a rule of inference is not an instance of knowing that. Let's call Ryle's moral 'the application problem.'

Ryle further argues that appealing to implicit propositional knowledge to get around the application problem is pointless. We can't explain why someone who has never been taught a principle of logic can nevertheless be a good reasoner just by saying that he has an implicit knowledge of the principles. Neither can we explain a skillful but theoretically unsophisticated chess player by saying that he has an implicit grasp of tactical maxims. The reason is that this "shuffle" fails to explain how someone can explicitly acknowledge truths and still "be a fool in his performance." (1945: 7-8). And knowing how to apply a rule or maxim comes, in his view, first: "we could not consider principles of method in theory unless we or others already intelligently applied them in practice. Acknowledging the maxims of a practice presupposes knowing how to perform it." The point is summarized by one of his best lines: "Rules, like birds, must live before they can be stuffed." (1945: 11).

What Ryle failed to see was that in addition to the application problem, there was also an adoption problem concerning rules of inference. He viewed Carroll's argument as a particular application of his own regress argument, and perhaps for that reason he failed to see that there is an important difference between basic rules of inference and the other rules he was considering.²⁰⁵ As we saw, Ryle treats the case of logic and the case of chess as if they were analogous. But it is not the case that to use a rule or principle of chess we need to apply the rule or principle to the rule or principle itself. There may be problems, as Ryle argues, in figuring out when to apply them or how to follow them, but the problem is not that we need to apply them to themselves in order to use them. And it is this last bit what makes the case of logic special and particularly challenging.

²⁰⁵ The fact that he did not discuss a particular application of a rule of inference and what it would take for someone to apply it may also have prevented him from noticing the differences.

Basic rules of inference provide the anti-intellectualist with a powerful argument since the main problem here is not that someone may fail to apply them but that, as Harry's case indicates, without the prior ability to infer it would be *impossible* to apply them. In this case propositional content is indeed inert. Therefore, they constitute a better example for Ryle's case against the intellectualists, since it is difficult – even in principle – to see how propositional content by itself can, explicitly or implicitly, guide their application. And they also support the second claim Ryle wants to make, namely, that the ability to infer comes first – the 'stuffing' occurs later.

That there is an adoption problem in addition to an application problem helps address the worry Stanley expresses in the following way:

But if this is the problem raised by the Lewis Carroll case [namely, that no matter how many propositions the agent appears to grasp, they may nevertheless fail to apply them], it simply cannot be solved by appeal to a non-propositional account of knowing how, or mere dispositions. If one has knowledge, whether propositional or non-propositional, how can there not be a possibility of failing to apply it when the time comes? The same point arises for mere dispositions. One can have all the dispositions one wants, but the exigencies of the situation may inhibit their manifestation (2011: 31).

If we see that the problem runs deeper and it is not just that we may fail to apply a basic rule of inference when the circumstances call for it, but that the very recognition of the circumstances for their application depends on a prior ability to infer (as it was argued in chapter 3), then the symmetry that Stanley is trying to establish between propositional knowledge and a disposition to infer or knowing how to perform an inference dissolves. The adoption problems calls into question the view that rules guide our basic deductive inferential performances, and with it the relevance of propositional content in the context of an inferential act.

This is exactly why setting up the problem *a la* Kripke is important. The problem is not that we may fail to apply propositional knowledge, but that it would be *impossible* to apply it if we didn't infer according to it already. Because of this, the case of basic logical rules or principles turns out to be a very strong case for the anti-intellectualists.

Would Stanley and Williamson's own brand of intellectualism escape an objection based on the adoption problem? In their view, knowing how to ϕ turns on knowing that w is a way of ϕ -ing, where w is introduced under a practical mode of presentation. Since they do acknowledge that there is a difference between knowing how and knowing that because the former introduces practical modes of presentation, it may be that their account can accommodate a case like Harry's.

Here is a summary of the view by Stanley himself:

According to Intellectualism about knowing how, knowing how to do something is equivalent to knowing in what way one could do it. So, for example, you know how to ride a bicycle if and only if you know in what way you could ride a bicycle. But you know in what way you could ride a bicycle if and only if you possess some propositional knowledge, viz. knowing, of a certain way w which is a way in which you could ride a bicycle, that w is a way in which you could ride a bicycle. This is essentially the view of knowing how advocated in Stanley and Williamson (2001). (2011b)²⁰⁶

Now, their proposal has generated many and varied objections, and there are many details that are omitted in this summary. But this basic idea will suffice to raise a worry in connection with Harry's case. Suppose Harry is as clueless as ever regarding UI. Saul is nearby and says: "all

²⁰⁶ Allow me once again to indulge myself in the habit of introducing quotes from Wittgenstein. He says: "Imagine a sequence of pictures. They shew how two people fence with rapiers according to such-and-such rules. A sequence of pictures can surely shew that. Here the picture refers to a reality. It cannot be said to shew *that* fencing is done like this, but *how* fencing is done. In another sense we can *say* that the pictures shew how one can get from this position into that in three movements. And now they also shew *that* one can get into that position in this way." (*RFM*: Part V, §51) [My emphasis]

ravens are black; this is a raven; therefore, it is black.” So, pointing to Saul, I say to Harry: “that is a way for you to perform a universal instantiation.” What would be required of Harry to realize that Saul’s way is a way for him to infer according to UI? It would seem that he would first have to formulate the UI rule and then see that that particular way, Saul’s way, is a way for him to perform a universal instantiation. But then we find ourselves in a similar situation to that of chapter 3 where examples were discussed: to recognize Saul’s way as a way for him to instantiate, Harry himself would have to perform an instantiation. The propositional knowledge introduced by means of a practical way of presentation would therefore seem to be insufficient to guide Harry’s inferential performances. (Note that things wouldn’t be any better if I stated the UI rule for Harry once again, as he would still have to perform an instantiation).

7.2.1 Basic Deductive Inferences as Knowing How

In the previous section it was argued that intellectualism cannot account for basic deductive inferential acts because the capacity or ability to perform basic inferences is presupposed in the application of rules of inference to particular inferential acts. Hence, the argument constitutes an objection to intellectualism for it identifies intelligent actions – and inferring according to rules of inference was supposed to be a paradigmatic case of an intelligent action – that cannot be considered intelligent in virtue of being guided by propositional knowledge of rules of inference.

More would need to be said, however, to establish that basic inferential acts are best explained in terms of knowing how. The reason is that performing a basic inferential act may not involve a state of knowledge at all, it may just be a habit or instinct, or a process resulting from a mechanism that cannot itself count as a state of knowledge. So, before considering whether

inferential acts can be considered instances of knowing how, we must have some account of knowing how. And here things get tricky because it appears that there are as many characterizations of knowing how as people discussing the notion.

Coming up with a characterization of what is it to know how to do something is, of course, an issue too broad and complex for us to deal with here. But we can, instead, consider some specific options. Let's start with Ryle's own approximation to the notion, paying special attention to its application to the case of inferential acts.

Ryle is usually taken to have held that to know how to Φ is to have the ability to Φ .²⁰⁷ Although it may be problematic to interpret Ryle as holding this in general, he seems to have held this for the case of inferential acts. In his view, knowing a rule of inference "is not possessing a bit of extra information but being able to perform an intelligent operation;" it is "knowing how to move from acknowledging some facts to acknowledging others;" and it is "realised in performances which conform to the rule, not in theoretical citations of it." (1945: 7)

There may be some hesitation on Ryle's part concerning the role of rules in inferential acts. On the one hand, he seems to say that we employ rules when performing inferences. On the other hand, he appears at times to suggest that we simply perform inferences that are later seen as conforming or not with rules that have been extracted from the practices and given formalization.

Consider again the analogy with chess. The chess player, whether smart or stupid, presumably has propositional knowledge of the rules of the game, even if their application ultimately turns out to be a matter of knowing how. And this suggests that the same holds for logic, since Ryle gives no indication of finding differences between the cases. Various passages appear to support

²⁰⁷ Hornsby (2012: 82), however, holds that Ryle had no definite account of the notion of knowing how and argues that, for Ryle, having an ability to do something is not sufficient for knowing how to do it.

this reading. For example, Ryle compares the observance of rules with wearing glasses: “we look through them but not at them;”(1945: 9) and he also claims: “to be acute and consistent in reasoning is certainly to apply rules of inference to the propositions considered. But the reasoner does not have both to consider propositions and to cast sidelong glances at a formula; he just considers the propositions efficiently. [...] The good chess player observes rules and tactical principles, but he does not think of them; he just plays according to them.” (1945: 14)

But some passages, on the other hand, can give the impression that the subject is not employing rules. For example, he claims: “when people can reason intelligently, logicians can then extract the nerve of a range of similar inferences and exhibit this nerve in a logicians’ formula [...] But arguing intelligently did not begin with Aristotle and does not after Aristotle require the separate acknowledgment of the truth or “validity” of the formula.” (1945: 7) And he also affirms: “A breach of a rule of logic is a fallacy; an observance of it is a valid inference. To speak of an inference as an observance or a fallacy as breach of a rule of logic is only a condensed way of saying that the author of the inference has made his inference in conformity with or in breach of a rule of inference.”

Ryle’s point seems to be, however, that although we are governed by rules in our performances, no acknowledgement of them, explicit or implicit, is required. He writes:

When a person knows how to do things of a certain sort (e.g., cook omelettes, design dresses or persuade juries), his performance *is in some way governed by principles, rules, canons, standards or criteria*. (For most purposes it does not matter which we say.) It is always possible in principle, if not in practice, to explain why he tends to succeed, that is, to state the reasons for his actions. It is tautology to say that there is a method in his cleverness. But his observance of rules, principles, etc., must, if it is there at all, be realised in his performance of his tasks. It need not (though it can) be also advertised in an extra performance of paying some internal or external lip-service to those rules or principles. (1945: 8-9) [My emphasis.]

The problem here is how to understand “in some way,” since Ryle does not say how rules are precisely supposed to govern deductive inferences. In the next section, we will consider one such possibility.

7.2.2 Hardwired Rules of Inference

There is a possibility we have not yet considered: even though we don’t consciously *follow* rules of inference when we perform inferential acts, we are *governed* by embodied or hardwired rules that are subconceptual and inaccessible to consciousness. A version of this view is held by Devitt (2006) as part of his argument that linguistic competence does not involve knowledge that. Somewhat reluctantly, as he would prefer not to talk of knowledge and simply say that linguistic competence is an ability or skill, he yields to the pressure of folk terminology and claims that “linguistic competence is mere knowledge-how.”²⁰⁸ (Devitt 2011c: 205) As it turns out, Devitt’s appeal to knowledge how sets in motion the ideas behind Stanley and Williamson (2001),²⁰⁹ who single him out (together with David Lewis) as a “representative case” of those holding that knowledge how is a feature of mental reality that is pertinent to psychological explanations. Recent papers and chapters (Devitt 2011a, 2011c and 2012; Stanley 2011a, 2011b, and 2013c)

²⁰⁸ Devitt thinks that knowledge how may sometimes involve propositional knowledge. He gives the example of chess: knowing how to play chess may include propositional knowledge of the rules of the game. In other cases, however, we may just have knowledge how, as in the case of knowing how to swim or ride a bike, which “prima facie does not involve any propositional knowledge at all.” (Devitt 2011a: 314) Linguistic competence belongs with this latter group.

²⁰⁹ Stanley (2011a: ix) recounts the story of how his joint paper with Williamson came about. It was their discussion of Devitt’s claim that knowledge of meaning is knowledge how (see Devitt 1981: 95-110 and 1997: 270-5) what led them to put together Stanley’s view that on all existing semantic accounts of embedded questions knowing how was a kind of propositional knowledge with Williamson’s view that knowing how to ϕ is to know, of some way w , that w is the way to ϕ .

indicate that neither party has yet succeeded in persuading the other, but our main interest here lies in Devitt's positive account of knowing how, and not so much in his arguments against Stanley and Williamson. The proposal here is to take Devitt, following Stanley and Williamson's suggestion, as a "representative case" of the hardwired rule view introduced in his discussion of knowledge how, though the morals will be quite general and, therefore, intended to cover (one hopes!) the non-representative cases of this view as well.

Devitt sees the "folk" distinction between knowledge that and knowledge how as more or less coinciding with the distinction between declarative knowledge and procedural knowledge. As he notes, the declarative-procedural distinction originated in AI, but it is commonly used in cognitive science (2006: 47). According to him, cognitive psychology and cognitive neuroscience understand propositional knowledge in terms of declarative knowledge (commonly associated with the ability to articulate such knowledge verbally²¹⁰), and distinguish it sharply from procedural knowledge, which is in turn understood as knowledge-how. Devitt characterizes the distinction as follows:

Psychologists describe the distinction, rather inadequately, along the following lines: where declarative knowledge is explicit, accessible to consciousness, and conceptual, procedural knowledge is implicit, inaccessible to consciousness, and subconceptual. (2009: 209)

Two other distinctions are related to this one. Firstly, the distinction between explicit and implicit memory, which corresponds to the declarative-procedural distinction. And secondly, the distinction between explicit and implicit learning, where the former is a top-down process that starts with declarative knowledge (typically exemplified by instructions in a user's manual), and

²¹⁰ Stanley and Krakauer (2013) argue that this association should be rejected. See below.

the latter is a bottom-up process that begins with observation and practice in order to pick up a skill.

In Devitt's view, however, the central point to understanding the distinction between declarative and procedural knowledge is to distinguish between "rules that govern by being represented and applied from ones that are simply embodied without being represented." (2006: 46) While psychologists agree that a conscious representation of what is known is implicated in declarative knowledge and, accordingly, that to have declarative knowledge of processing rules a subject must possess a conscious representation of them, it is not clear that procedural knowledge requires the representation of the rules. Devitt points out that, even though there is no consensus on this issue, the psychologists who hold that procedural knowledge involves representing processing rules claim that they are different representations from those implicated in declarative knowledge.

One of the many examples he provides in favor of drawing the distinction between represented rules and embodied rules is that of the kingfisher:

The kingfisher catches fish by diving into the water [...] in diving, it must make suitable allowance for the refraction of light [...] it would be absurd to suggest that the kingfisher goes through the following process: it represents (tacitly knows) that for μ water-to-air refraction is such and such; it represents that the angle ϕ is so and so; it represents that $\mu = \sin \phi / \sin \theta$; it uses this information to calculate the angle of θ . Applying this information it dives into the water at the appropriate angle. It is not plausible to suppose that the kingfisher represents any of these facts about refraction and angles. No more does it represent whatever rules do govern its fishing: "mechanisms that perform certain tasks need not embody theories of those tasks" (Schank 1980: 35). Those mechanisms, reflecting information about water-to-air refraction, are simply built into the kingfisher "by nature" (2006: 49)

Interestingly enough, Devitt maintains that inferring itself involves embodied rules. He claims:

Finally, consider the most cognitive skill of all, the skill of thinking, of inferring one thought from another. We can, of course, learn “laws of thought” – that is what we do in logic classes – but most of us know hardly anything about these rules. Most thinking in most people surely does not involve applying representations of these rules, representations involved in higher-level thoughts about the rule-governed thoughts; and *all* thinking could not, as Lewis Carroll showed. (p. 50)

There is a methodological moral, for Devitt, to be drawn concerning processing rules, which he calls “Pylyshyn’s Razor,” and according to which “representations are not to be multiplied beyond necessity.” The idea is that we should only introduce represented rules in those cases in which representations have an explanatory work to do, and by this he means causal work. He says that the rationale for his methodological approach is mainly epistemological: failure to introduce representations when they are indeed needed would probably yield inadequate explanations, while their introduction when there are no such representations can be hardly defeated by evidence “because with enough representations almost any behavior can be explained.” (52)

Now, the question becomes whether basic rules like UI can be thought of as being hardwired, or embodied, in us. It is clear that, on this view, the fact that Harry is unable to adopt the UI rule will turn out to be hardly surprising. And indeed, one might even see the adoption problem as giving support to a hardwired proposal (much to Kripke’s own discomfort). If we still want to help Harry out, we should be thinking more along the lines of brain surgery than telling him anything. For all we know, basic deductive hardwired rules could be induced in a less dramatic way, especially if – as Devitt claims (2006: 47) – relevant modifications to the wetware as a result of experiences are indeed possible for the case of basic rules. Yet, if we assume that Harry has been extensively exposed to more or less the same experiences than most people, then there is

hardly any hope for him. What is clear is that stating the rule for him would be completely pointless.

Let's not, however, get ahead of ourselves. One issue to consider is whether inferring can be counted as a skill. In chapter 2, we compared basic logical principles with other candidates for *a priori* knowledge, like mathematical or conceptual statements, and argued that the impossibility of adopting them made them especial. But the issue was raised of whether inferring in accordance with UI could be more akin to basic skills, like boiling an egg or even riding a bike. So the question turned then on whether *skills* could be up for adoption: what would happen if we were given instructions for very basic skills, like the ones just mentioned? We left this question open, on the idea that this issue would be better discussed together with the distinction between know how and know that. It is time now to address the issue.

Despite Stanley's efforts to the contrary, there seems to be a wide consensus among philosophers, cognitive psychologists and cognitive neuroscientists that skills don't require propositional knowledge and belong in the category of procedural knowledge. Now, there is also considerable agreement that even when propositional knowledge is present in the acquisition of a skill, it becomes superfluous, or even detrimental, once the skill is in place (Dreyfus 2005; Dreyfus and Dreyfus 1980). Psychological studies suggest that skill learning is in many cases implicit, that is, a bottom-up process. Skills are picked up by observation and practice, without our relying on explicit propositional knowledge of the principles or rules involved in the process (see Reber 2003 and Sun, Merrill, and Peterson 2001).²¹¹

²¹¹ Devitt says: "Even if representations of some of these rules can play a role in the acquisition and improvement of the skills, it is not plausible to suppose that those rules must be represented in a person once she has achieved competence ..." (2006: 49-50)

The standard view on skills has recently come under pressure (see Stanley 2011 and Stanley and Krakauer 2013). Stanley and Krakauer argue that the manifestation of a skill is “guided by knowledge of facts about that activity.” (3) Knowing what to do to initiate an action is an example of the kind of knowledge that, they claim, is needed for skilled intentional actions: “in the case of virtually any activity ϕ , having skill at ϕ -ing requires knowing what to do to initiate actions of ϕ -ing” and “such knowledge is propositional knowledge.” (4) Agents may not be able to verbalize their knowledge, but they argue that the connection between propositional knowledge and the linguistic ability to articulate it should be rejected. Skills have, according to Stanley and Krakauer, two components, an “ever accumulating propositional knowledge about that activity” (9) and a performance (or procedural) aspect that gets better with practice and concerns the selection and motor acuity of the different actions that may be involved in the skill.

Without taking sides, let me say the following. Even if we agree with the view that many skills are picked up from observation and practice without the need of declarative knowledge, that does not mean that a skill could not be learnt by appealing to declarative knowledge. Consider the case of boiling an egg. Someone who has never cooked anything in her life could just follow a set of instructions that carefully specify the temperature of the water, the time, and so on, and eventually produce a boiled egg. Even learning a more complex skill in this way, like riding a bike, seems to be in principle possible. Most of us have learned to ride a bike by observation and practice (maybe with a little help from someone holding the bike at first). It is likely that, if asked, we would struggle with a verbalization of how we did it. But the relevant question for us is the following: could we learn to ride a bike if we were just given a set of directions? I don’t see why not. I can imagine the instructions saying things like “To maintain equilibrium keep pedaling at an even pace,” or “when equilibrium is lost and the bike is leaning towards one side, the rider

should lean towards the other side to compensate,” “for a full stop press the brakes and put at least one foot on the ground.” Of course, a set of directions such as this won’t make us proficient, won’t give us what Stanley and Krakauer call ‘motor acuity,’ but that is a different issue. With practice and patience alone we could become so. It is conceivable that someone who has never seen anybody else ride a bike would be capable of picking it up just from a set of directions. In contrast, we won’t get Harry to infer in accordance with UI from a set of directions.

If it is indeed possible to learn a skill without the need of declarative knowledge, this would suggest that declarative knowledge is *not essential* to knowledge-how. In contrast, if inferring according to UI were a skill, it would be a special kind of skill that *cannot* be acquired by means of declarative knowledge. It would then be a case in which declarative knowledge is *superfluous* for knowing how.

One reason for thinking that inferring according to basic rules of inference cannot be a skill, but at most a capacity or ability, is that there is ample evidence that we are indeed notoriously bad at performing inferences. Talk of skill seems to be associated with some kind of excellence, and we wouldn’t say that someone who is capable of (say) riding horses but is a mediocre rider possesses a skill. So, it would seem that, although some of us may be skilled at inferring, most of us are not.

This is connected with a separate and more important issue. A discussion of embodied or hardwired rules appears to make more sense, for instance, in the case of linguistic competence, as we are remarkably good at observing the rules postulated by the universal grammar. In contrast, we seem to be, as we said, very bad at performing basic deductive inferences. And the reasons for our poor performances seem to be quite varied. We are, for example, highly biased to content. Consider the following inference:

Premise 1: All living things need water

Premise 2: Roses need water

Therefore, Roses are living things

Cognitive psychologists have found, while looking into belief bias effects, that roughly 70 percent of university students presented with this problem incorrectly respond that the conclusion is valid (Markovits and Nantel 1989; Sá, West, and Stanovich 1999; Stanovich and West 1998c). It seems evident that the believability of the conclusion interferes with the appraisal of logical validity since virtually all of the same students correctly pronounce the following structurally similar syllogism to be invalid:

Premise 1: All insects need oxygen

Premise 2: Mice need oxygen

Therefore, Mice are insects

Moreover, there are several factors that influence performances and the evaluation of inferences which appear to be quite independent from the abilities involved in performing inferences. For instance, other studies have found that a higher degree of criticism among individuals of belief-inconsistent evidence over belief-consistent evidence is unrelated to cognitive ability (Klaczynski 1997; see also Klaczynski and Gordon 1996; Klaczynski et al. 1997; Klaczynski and Lavallee 2005; Klaczynski and Robinson 2000), and also that a subject's idiosyncratic measure of dogmatism and absolutism can be used to predict variance in argument appraisal even after cognitive ability is partialled out. Stanovich (2008) claims that "variation in cognitive ability indexes individual differences in the efficiency of processing at the algorithmic

level. In contrast, thinking dispositions index individual differences at the intentional level.” His point is that thinking dispositions have to do with the epistemic goals of the individuals: while someone who agrees with “People should always take into consideration evidence that goes against their beliefs” values being accurate more than preserving her beliefs, someone who disagrees with such statement signals that she values preserving her beliefs over accuracy. These thinking dispositions offer, hence, information about epistemic goals at the intentional level, and Stanovich’s claim is that an adequate account of individual variances in reasoning needs to take into consideration not only differences at the algorithmic level (computational capacity), but also the intentional level, the thinking dispositions that reveal the epistemic goals of the individuals.²¹²

Moreover, the two more influential theories of human reasoning, the mental logic theory (Rips) and the mental models theory, presuppose that the rules or models are represented by the thinkers. But neither of these theories seems to have gained substantial consensus. In fact, quite the opposite appears to be the case.²¹³ And cognitive psychologists seem to be either trying to add new elements to theories of deductive reasoning that are extrinsic to the cognitive reasoning abilities of the thinkers, such as Stanovich’s “thinking dispositions,” or appealing to a more basic

²¹² Stanovich (2008) says: “The instructions for many tasks which require reasoning in the face of belief bias (e.g., Baron 1995; Evans, Newstead, Allen, and Pollard 1994; Handley et al. 2004; Oakhill, Johnson-Laird, and Garnham 1989; Stanovich and West 1997) dictate that prior belief be totally discounted in evaluating the argument. But individuals may differ in their willingness as well as their ability to adapt to such instructions. Some individuals may put a low priority on allocating computational capacity to evaluate the argument. Instead, for them, capacity is engaged to assess whether the conclusion is compatible with prior beliefs (Evans, Barston, and Pollard 1983; Evans et al. 1994).”

²¹³ Consider Stenning and Van Lambalgen’s assessment of the situation: “From the position of being absolutely central in the cognitive revolution, which was founded on conceptions of reasoning, computation and the analysis of language, the psychology of deduction has gone to being the deadbeat of cognitive psychology, pursued in a ghetto, surrounded by widespread scepticism as to whether human reasoning really happens outside the academy. “Isn’t what we *really* do decision?” we increasingly often hear. Many eminent psychology departments do not teach courses on reasoning. Imagine such a psychology department (or indeed any psychology department) not teaching any courses on perception. Even where they do teach reasoning they are more likely to be focused on analogical reasoning, thought of as a kind of reasoning at the opposite end of some dimension of certainty from deduction.” (Adler and Rips 2008: 223).

notion of interpretation (Stenning and Van Lambalgen) to account for individual differences, or changing the picture altogether and claiming that what goes on in reasoning cannot be modeled deductively.

Cognitive science doesn't have, at least at the moment, many answers for us regarding inferential competence. We have no clear idea of what is supposed to be embodied. It is not evident that it should be only one "mechanism." We don't know either whether rules, mental models, or something else is involved, and even though cognitive scientists seem to think that representations are implicated, they may well be wrong given the uncertainties that surround the whole issue of the psychology of deductive reasoning.

There is an additional problem in thinking that, by claiming that the rules of inference are hardwired, we avoid the problem of presupposing the ability to infer according to them. One of the examples offered by Devitt of a mechanism that operates on embodied rules is a calculator. The calculator does not 'represent' the rules. It is simply built in such a way that it operates on them without having to represent their content. The numbers, however, introduced as the inputs for the operations are represented. Consider now, in a similar way, a mechanism that operates on UI without representing it. But even in that case, whatever specialized procedure the mechanism embodies will have to be general. And wouldn't a UI inference be needed for an input to be recognized as an input? But if that is the case, whatever the advantage introduced in hardwired rules would be lost.

All this suggests that it is highly doubtful that something akin to basic rules of inference are embodied in us. It may be a good idea here to apply Devitt's methodological razor not only to representations. Why stop there? After all, the same methodological concerns would apply to rules of inference as well. And it is unclear at this point whether they play an explanatory role. If

what we have been saying is correct, there are reasons for thinking that that is not the case. If a recommendation were to be made, I would favor a ‘healthy agnosticism’ with respect to them.

7.2.3 Basic Inferring without Rules

Up to this point we have tried to make sense of the idea that inferential acts which accord with basic rules of inference are somehow informed by the rules themselves. It is time now to give serious thought to the possibility that, at least in the case of some basic inferences, we simply don’t rely on rules of inference.

Can we make sense of this possibility? It is by no means clear. Inferences are supposed to be a paradigmatic case of rule following, and it is indeed hard to see what is left once rules are taken away from them. If rules don’t play either an implicit or explicit role in basic inferential acts, the problem is how to understand the relation between premise(s) and conclusion. The whole point of appealing to rules was precisely that their content was supposed to guide the inferring thinker and secure an appropriate relation: it was because the thinker was somehow riding on a rule that the conclusion could be seen as following from the premise(s) in the way specified by the rule. That an inference accords with a rule is, on this view, anything but arbitrary: it accords with a rule because it was determined by a rule. Once we come to suspect that the inference ticket model of rules of inference – where rules are what ensures that a conclusion is reached in the ‘right’ way – leaves us exactly where we where, that to apply the rule we would need another ticket of the very same kind, we need another way of understanding the relation between premise(s) and conclusion. The danger is that if there is no such role for rules to play, the passage

from premise(s) to conclusion would seem to be, in Kripkean language, ‘a leap in the dark’ from the thinker’s perspective. And yet it does not feel that way at all. The question is: why?

The view I want to air here contemplates the possibility that basic inferential acts are to be understood as a primitive skill or ability that sets in motion regularities or patterns of use – where judgments are made (or sometimes simply expressed by means of an action) as a result of other judgments – and which constitute the basis of our *talk* of inferential transitions as a form of rule following.

I should say at the outset that I don’t have a settled view on whether this picture can ultimately be made to work, and so the following considerations will have to be sketchy and should be taken to be provisory, representing more a point in which my thought has come to rest at the present time, than a final solution to the problems that have been posed.

We need to rethink what kind of relation premise(s) and conclusion can plausibly have in an inferential act. If rules cannot provide us with the reasons for establishing a connection, we have to look somewhere else. And there are not many other places to look.

Let’s go back to the beginning. My nephew tells me that all the animals in *Madagascar* talk. I believe him but still ask: “does Alex the lion talk?” (Remember, I am still asleep). His impatient reply is *immediate*: “I told you that all the animals in the movie talk!” We have left the zoo and are now at a party. I announce that Maria had told me that if she came to the party, she would sing for us. A friend says that she has already seen Maria, and I *immediately* add, “great, I was hoping to hear her sing.” Someone else intervenes and asks why we believe Maria will sing. What’s the most likely reaction in such scenario? Discounting walking away without saying a word, the most plausible answer is that on a first attempt we will repeat the premises. Only when it is clear that

our premises are understood and not in doubt we might try to appeal to the nothion of validity, or to the MPP rule of inference.

These reactions may suggest that the premises themselves are all the reasons we have for performing an inference, at least at the very basic level. Perhaps Kripke has in mind something along these lines when he says:

I mean look, what would my own standpoint be? What's the naive standpoint about all of this? It's that, well, from p and q , the conjunction follows, independently of what logic you accept. Accepting logic has nothing to do with it. If someone doesn't accept my logic, that doesn't matter, ' p and q ' will follow. He says p and he says q . If he doesn't infer it, that's so much the worse off for him, he hasn't inferred something which really follows. But if you don't like that standpoint, I mean, that would be the old sort of reactionary standpoint. If you don't like that standpoint, it's no good to say, well, from p and q , ' p and q ' provided you accept a standard logic, or if you accept this rule of inference. There is nothing else that you can accept which is going to get you to this conclusion if not already.

In chapter five, we found Kripke defending a view based on the notion of intuition, where basic principles and rules of logic were intuited as self-evident. But we anticipated then that it is possible to track another line of thought in his seminars, particularly in his Princeton Seminar. The influence of Part I of the *Remarks on the Foundations of Mathematics* is evident,²¹⁴ though he rarely mentions Wittgenstein. [We should say, however, that he has since then become more of a hard core Platonist and it is almost surprised by some of the things he said.]

The more Wittgensteinian line of thought emphasizes the importance of being in the *habit* of inferring in accordance with basic logical principles. The passages hardly constitute a view, but they are suggestive. For example:

²¹⁴ Recall that the seminar was intended to be on the material that later became *WRPL*.

If someone did not already have the *habit* of inferring an instance of a universal statement from the universal statement itself, no amount of telling him that it's a law of logic that all universal statements imply each of their instances is going to get one any further. That's the point in a nutshell. (Princeton Seminar)

And, in connection with the possibility of giving someone like Harry a rule instead of a principle, he says:

... if he is the kind of person who was not in the *habit* of accepting inferences of this pattern already, he would not get anything from your rule either. [...] Of course, in a sense, the [inaudible] of all of us consists in the adoption of a rule. But what does that mean? Just that we *conform* to the rule. But it doesn't mean that God told us to obey this rule and then we said you know, we will listen and obey or something like that, because had he done so we couldn't have followed his instructions without being in this particular habit of making this particular kind of inference already.

Later on, discussing the perverse instantiation example, he says:

... we were trying to get him to exclude inference patterns of this form, but he, because he had this *habit* already of using a non-standard logic, actually concluded on the basis of what we told him, that this inference is valid, not fallacious.

Kripke does not make explicit what he means exactly by 'habits,' but he does make clear that it is *not* a statement of a basic principle or rule we have in mind when we reason what helps us perform an inference. Instead, in a more Wittgensteinian spirit – "... it is not a kind of *seeing* on our part; it is our *acting*, which lies at the bottom of the language-game" (*On Certainty* § 204) –, he appears to think that habits lie beneath basic rules and principles. Because we already make such inferences, because we are in the *habit* of using them, abstract principles about them can be stated and we can see that they hold.

... at least some of the laws of logic have a special character pointed out by Lewis Carroll, that they cannot be regarded as assumptions, basic beliefs, ingrained hypothesis, such that because we believe them we do certain things, because such a belief, whether we have it or not, doesn't help us .

[...]

Eventually one just can't give oneself a belief or a premise or something, one just does draw certain conclusions.

On the contrary, it's because we make all of these inferences that we can state them as an abstract principle and we can see that it represents a general account of what we do. Kripke seems to go a step further when he says (following Wittgenstein):

Maybe that is what to accept a logic is. It is to infer in accordance with a certain pattern. But that, so to speak, is not an intellectual process like the adoption of a hypothesis. Well, it might be an intellectual process of some force that we do *conform* to this pattern. Obviously we see that certain things follow from certain others. Rather than try and think of this as mysterious, that if all things are black the Eiffel tower has got to be black. Recognizing a principle of logic here seems to me to be an unnecessary shuffle, to use a Wittgensteinian phrase. A piece of extra baggage.

So, not only are the laws of logic not helpful when understood as beliefs, assumptions, or hypotheses, but also he seems to think that their formulation is superfluous. [This is the kind of passage that he would find excessive nowadays] We *conform* to them, but do not go by them.

Let me just propose the following, without claiming that this was what Kripke had in mind. It is a primitive fact about us that we expect some things to follow from others. That's simply something we do. It is a brute fact that applies to humans and animals alike. At the most basic level is our *logica utens*. We take a conclusion to follow from some premise(s), and our *sole* reason

for that is given by the premise(s). We act on them in certain ways and repeat the same reactions when we encounter similar situations or information.

One of the consequences of the view seems to be that at the most basic levels we don't "entertain any logical theory." We do have the intuition, nevertheless, that Matías knows something that Harry doesn't. He can produce inferences and correct others in a basic and unsophisticated way. And it seems that the most natural way to account for this is to appeal to the notion of knowing how. That's not, however, all we can do. We can also identify patterns of use by reflecting on what we do, and this is also a natural part of our inferential practice. Reflection and the ability to identify patterns of use make the formulation of general principles or rules like UI or MPP possible. We can then *say* that our particular inferential acts conform or do not conform to them. But that doesn't mean that general rules operate at the level of the inferential act itself.

How are our attributions of knowing how to infer to be understood? When are we in position to say that someone knows how to infer? It seems to me that we don't need to suppose that such attributions can only be made from the *logica docens* point of view. Inside the *logica utens* itself, the recognition of our inferences by others could be thought as an attribution of knowing how. Once inferential patterns are formed, we also have the capacity to reflect on them (maybe it is at this point that we part company with animals). At a very primitive level, such reflection could be expressed by criticism of inferences that others or we make. This is already outside the inferential practice as it implies reflection about it. It is already part of the *logica docens*, though still at a very primitive stage. More sophisticated attributions will be possible as reflection on the practice progresses.

References

[For all cases where reprints are listed, references are to the reprint]

- Adams, E. W. (1975). *The Logic of Conditionals: An Application of Probability to Deductive Logic*. Reidel.
- Adler, J. E. (2003). The Revisability Paradox. *Philosophical Forum* 34 (3-4): 383–390.
- Adler, J. E. & Rips, L. J. (eds.) (2008). *Reasoning: Studies of Human Inference and its Foundations*.
Cambridge University Press.
- Berry, G. (1968). Logic with Platonism. *Synthese* 19 (1-2): 215 - 249.
- Bealer, G. (1998). Intuition and the Autonomy of Philosophy. In DePaul and Ramsey 1998: 201-40.
- Bealer, G. (1999). A Theory of the A Priori. In Tomberlin (1999).
- Belnap, N. (1962). Tonk, Plonk and Plink. *Analysis* 22 (6): 130-134.
- Benacerraf, P. (1965). What numbers could not be. *Philosophical Review* 74 (1): 47-73.
- Benacerraf, P. (1973). Mathematical truth. *Journal of Philosophy* 70 (19): 661-679.
- Bengson, J. & Moffett, M. A. (eds.) (2011). *Knowing How: Essays on Knowledge, Mind, and Action*.
Oxford University Press.
- Besson, C. (2009). Logical knowledge and Gettier cases. *Philosophical Quarterly* 59 (234): 1-19.
- Birkhoff, G. and Neumann, J. von (1936). The Logic of Quantum Mechanics. *Annals of Mathematics* 37: 823-843.
- Black, M. (1970). The Justification of Logical Axioms. In *Margins of Precision*. Cornell University Press.
- Blackburn, S. (1995). Practical Tortoise Raising. *Mind* 104 (416): 696-711.

- Block, N. (2000). Semantics, Conceptual Role. In Craig, E. (ed.), *Routledge Encyclopedia of Philosophy*, Routledge.
- Bobzien, S. (1997). The Stoics on Hypotheses and Hypothetical Arguments. *Phronesis* 42 (3): 299-312.
- Bobzien, S. (2002). The Development of Modus Ponens in Antiquity: From Aristotle to the 2nd Century AD. *Phronesis* 47 (4): 359-394.
- Boghossian, P. (1989). The Rule Following Considerations. *Mind* 98 (392): 507-549. Reprinted in Boghossian 2008b: 9-50.
- Boghossian, P. (1996). Analyticity Reconsidered. *Noûs* 30 (3): 360-391. Reprinted in Boghossian 2008b: 195-224.
- Boghossian, P. and Peacocke, C. (eds.) (2000). *New Essays on the A Priori*. Oxford University Press.
- Boghossian, P. (2000). Knowledge of Logic. In Boghossian and Peacocke 2000: 229-254.
- Boghossian, P. (2001a). How are Objective Epistemic Reasons Possible? *Philosophical Studies* 106 (1-2): 340-380. Reprinted in Boghossian 2008b: 235-266.
- Boghossian, P. (2001b). Inference and Insight. *Philosophy and Phenomenological Research* 63 (3): 633-640.
- Boghossian, P. (2003). Blind Reasoning. *Proceedings of the Aristotelian Society, Supplementary Volume* 77 (1): 225-248. Reprinted in Boghossian 2008b: 267-287.
- Boghossian, P. (2004). Epistemic Analyticity: A Defense. *Grazer Philosophische Studien* 66 (1): 15-35. Reprinted in Boghossian 2008b: 225-233.
- Boghossian, P. (2005). Rules, Meaning, and Intention. *Philosophical Studies* 124 (2): 185-197.
- Boghossian, P. (2008a). Epistemic Rules. *Journal of Philosophy* 105 (9): 472-500. Reprinted in Boghossian 2008b: 109-136.

- Boghossian, P. (2008b). *Content and Justification*. Oxford University Press.
- Boghossian, P. (2011). Williamson on the A Priori and the Analytic. *Philosophy and Phenomenological Research* 82 (2): 488-497.
- Boghossian, P. (2012). Inferentialism and the Epistemology of Logic: Reflections on Casalegno and Williamson. *Dialectica* 66 (2): 221-236.
- Boghossian, P. (2014). What Is Inference? *Philosophical Studies* 169: 1-18.
- Bonjour, L. (1998). *In Defense of Pure Reason*. Cambridge University Press.
- BonJour, L. (2001). Precis of *In Defense of Pure Reason*. *Philosophy and Phenomenological Research* 63 (3): 625–31.
- Borges, J. L. (1964). “The Avatars of the Tortoise.” In *Labyrinths: Selected Stories and Other Writings*, Yates, D. A. and Irby J. A. (eds), New Directions Publishing Co.: 202-208.
- Brandom, R. (1994). *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press.
- Brandom, R. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Harvard University Press.
- Broome, J. (2014). What Is Inference? *Philosophical Studies* 169: 19-25.
- Brown, D. G. (1954). What the Tortoise Taught Us. *Mind* 63 (250): 170-179.
- Burge, T. (1993). Content Preservation. *Philosophical Review* 102 (4): 457-488.
- Cappelen, L. (2012). *Philosophy without Intuitions*. Oxford University Press.
- Carey, S. (2009). Where Our Number Concepts Come From. *Journal of Philosophy* 106 (4): 220-254.
- Carnap, R. (1937). *The Logical Syntax of Language*. Harcourt Brace.
- Carnap, R. (1950). Empiricism, Semantics and Ontology. *Revue Internationale de Philosophie* 4: 20-40.

- Carroll, L. (1895). What the Tortoise Said to Achilles. *Mind* 4 (14): 278-280.
- Carruthers, P. (2004). On Being Simple Minded. *American Philosophical Quarterly* 41 (3): 205-220.
- Carruthers, P. (2009a). How We Know Our Own Minds: the Relationship Between Mind-reading and Metacognition. *Behavioral and Brain Sciences* 32(2): 1-62.
- Carruthers, P. (2009b). Cartesian Epistemology: Is the Theory of the Self-Transparent Mind Innate? *Journal of Consciousness Studies* 15 (4): 28-53.
- Cassam, Q. (2000). Rationalism, Empiricism, and the A Priori. In Boghossian and Peacocke 2000: 43-64.
- Chalmers, D. (2011). Revisability and Conceptual Change in “Two Dogmas of Empiricism.” *Journal of Philosophy* 108 (8): 387-415.
- Church, A. (1974). Outline of a revised formulation of the logic of sense and denotation (part II). *Noûs* 8 (2): 135-156.
- Coffa, A. (1991). *The Semantic Tradition from Kant to Carnap*. Cambridge University Press.
- Colyvan, M. (2006). Naturalism and the Paradox of Revisability. *Pacific Philosophical Quarterly* 87 (1): 1–11.
- Creath, R. (1987). The Initial Reception of Carnap’s Doctrine of Analyticity. *Noûs* 21 (4): 477-499.
- Dehaene, S. (1997). *The number sense*. Oxford University Press.
- Davidson, D. (1982). Rational Animals. *Dialectica* 36 (4): 317-28. Reprinted in Davidson 2001: 95-106.
- Davidson, D. (2001). *Subjective, Interobjective, Objective*. Oxford University Press.
- DePaul, M. and Ramsey, W. (eds.) (1998). *Rethinking Intuition*. Rowman and Littlefield.

- Devitt, M. (1998). Naturalism and the A Priori. *Philosophical Studies* 92 (1/2): 45-65. Reprinted in Devitt 2010.
- Devitt, M. (2005). There is No A Priori. In Steup and Sosa 2005: 105-115.
- Devitt, M. (2006a). *Ignorance of Language*. Oxford University Press.
- Devitt, M. (2006b). Intuitions. Reprinted in Devitt 2010: 293-302.
- Devitt, M. (2010). *Putting Metaphysics First: Essays on Metaphysics and Epistemology*. Oxford University Press.
- Devitt, M. (2011a). Linguistic Knowledge. In Bengson and Moffett 2011.
- Devitt, M. (2011b). No Place for the A Priori. In *What Place for the A Priori?* Michael J. Shaffer and Michael L. Veber, eds. Chicago and La Salle: Open Court Publishing Company (2011): 9-32. Reprinted in Devitt (2010).
- Devitt, M. (2011c). Methodology and the Nature of Knowing How. *Journal of Philosophy* 108 (4): 205-218.
- Devitt, M. (2012). Linguistic Knowledge. *Philosophical Inquiry* 36 (1): 314.
- Descartes, R. (1641). *Sixth Set of Replies*. In *The Philosophical Writings of Descartes*, Vol. II, Cottingham, J., Stoothoff, R., and Murdoch, D. (eds. and trans.). Cambridge University Press, 1984.
- Descartes, R. (1644). *The Principles of Philosophy*. In *The Philosophical Works of Descartes*, Vol. 1. Haldane, E. and Ross, G.R.T. (eds. and trans.). Cambridge University Press, 1969.
- Dodgson, C. L. (1977). *Lewis Carroll's Symbolic Logic*. Bartley III, W. W. (ed.). Clarkson Potter.
- Dogramaci, S. (2010). Knowledge of Validity. *Noûs* 44 (3): 403-432.
- Dogramaci, S. (2012). Apriority. In Russell, G. and Graff Fara, D. (eds.), *The Routledge Companion to Philosophy of Language*. Routledge.

- Dogramaci, S. (2013). Intuitions for Inferences. *Philosophical Studies* 165 (2): 371-399.
- Dummett, M. (1973a). The Philosophical Basis of Intuitionistic Logic. In Dummett 1978: 215-247.
- Dummett, M. (1973b). The Justification of Deduction. In Dummett 1978: 290-318.
- Dummett, M. (1973c). The Significance of Quine's Indeterminacy Thesis. In Dummett 1978: 375-419.
- Dummett, M. (1973d). *Frege: Philosophy of Language*. Duckworth.
- Dummett, M., (1976). Is Logic Empirical? In Dummett 1978: 269-289.
- Dummett, M. (1978). *Truth and Other Enigmas*. Harvard University Press.
- Dummett, M. (1991). *The Logical Basis of Metaphysics*. Harvard University Press.
- Dreyfus, S. E. and Dreyfus, H. L. (1980). *A Five-Stage Model of the Mental Activities Involved in Directed Skill Acquisition*. University of California Berkeley Operations Research Center.
- Dreyfus, H. L. (2005). Overcoming the myth of the mental: how philosophers can profit from the phenomenology of everyday expertise. *Proceedings and Addresses of the American Philosophical Association* 79 (2): 47-65.
- Engel, P. (2005). Logical Reasons. *Philosophical Explorations* 8 (1): 21-35.
- Euclid (1956). *Euclid's Elements*, Sir Thomas Little Heath (trans.). New York, Dover.
- Field, H. (1996). The A Prioricity of Logic. *Proceedings of the Aristotelian Society* 96: 359-379.
- Field, H. (1998). Epistemological Nonfactualism and the A Prioricity of Logic. *Philosophical Studies* 92 (1/2): 1-24.
- Field, H. (2000). Apriority as an Evaluative Notion. In Boghossian and Peacocke 2000: 117-149.
- Field, H. (2005). Recent Debates on the A Priori. In *Oxford Studies in Epistemology*, Szabo Gendler, T. and Hawthorne, J. (eds), Oxford University Press, Chapter 3.

- Field, H. (2009). What is the Normative Role of Logic? *Aristotelian Society Supplementary Volume*, 83 (1): 251-268.
- Finkelstein, D. (1964). Matter, Space, and Logic. *Boston Studies in the Philosophy of Science V*, Cohen and Wartofsky (eds.), Kluwer.
- Fodor, J. (1987). *Psychosemantics: The Problem of Meaning in Philosophy of Mind*. MIT Press.
- Fodor, J. and Lepore, E. (2006). Analyticity Again. In Michael Devitt & Richard Hanley (eds.), *The Blackwell Guide to the Philosophy of Language*. Blackwell.
- Fodor, J. (2008). *LOT 2, The Language of Thought Revisited*. Oxford University Press.
- Frankfurt, H. (1977). Descartes on the Creation of the Eternal Truths. *Philosophical Review* 86 (1): 36-57.
- Frege, G. (1884). *Die Grundlagen der Arithmetik*, Breslau: Verlag von Wilhelm Koebner. Translated by J. L. Austin as *The Foundations of Arithmetic*. Blackwell, 1950.
- Frege, G., (1893). *Grundgesetze der Arithmetik*, Vol. 1. In *The Frege Reader*, Beaney, M. (ed. and translator).
- Frege, G. (1899-1906). "On Euclidean Geometry." In Frege (1979): 167-169.
- Frege, G. (1979). *Posthumous Writings*. Edited by Hans Hermes, Friedrich Kambartel, and Friedrich Kaulbach, Chicago, The University of Chicago Press.
- Friedman, M. (1999). *Reconsidering Logical Positivism*. Cambridge University Press.
- Gallistel, C. R. (1990). *The organization of learning*. MIT Press.
- Gardner, M., (1971). Is Quantum Logic Really Logic? *Philosophy of Science* 38 (4): 508-529.
- Geach, P. (1965). Assertion. *Philosophical Review*, 74 (4). Reprinted in Geach 1972: 254-269.
- Geach, P. (1972). *Logic Matters*. Blackwell.
- Geach, P. (1973). Omnipotence. *Philosophy* 48 (183): 7-20.

- Ginet, C. (1975). *Knowledge, Perception, and Memory*. Reidel Publishing Company.
- Goldman, A. (1992). A Priori Warrant and Naturalistic Epistemology. In Tomberlin (1999).
- Goodman, N. (1955). *Fact, Fiction, and Forecast*. Harvard University Press.
- Johnson-Laird, P. (1983). *Mental Models*. Harvard University Press.
- Johnson-Laird, P. and Byrne, R. (1991). *Deduction*. Hillsdale, NJ, Lawrence Erlbaum.
- Hale, B. and Wright, C. (2000). Implicit Definition and the A Priori. In Boghossian and Peacocke 2000: 286-319.
- Hale, B. (2002). "Basic Logical Knowledge." In *Logic, Thought, and Language*, O'Hear A. (ed.), Cambridge University Press: 279-304.
- Hanna, R. (2006a). Rationality and the Ethics of Logic. *Journal of Philosophy* 103 (2): 67-100.
- Hanna, R. (2006b). *Rationality and Logic*. MIT Press.
- Harman, G. (1973). *Thought*. Princeton University Press.
- Harman, G. (1984). Logic and reasoning. *Synthese* 60 (1): 107-127.
- Harman, G. (1986). *Change in View: Principles of Reasoning*. MIT Press.
- Harman, G. (1995). "Rationality. In *Thinking: an Invitation to Cognitive Science*, Smith, E. and Osherson, D. (eds.), Volume 3. MIT Press: 175-211.
- Harman, G. (1996). Analyticity Regained? *Noûs* 30 (2): 392-400.
- Harman, G. (2002). "Internal Critique: A Logic is not a Theory of Reasoning and a Theory of Reasoning is not a Logic." In *Handbook of the Logic of Argument and Inference: The Turn Towards the Practical*, Vol I. Gabbay, D. M., Johnson, R. H., Ohlbach, H.J., and Woods, J. (eds.), Elsevier Science: 171-186.

- Harman, G. (2003). The Future of the A Priori. *Philosophy in America at the Turn of the Century, APA Centennial Supplement to Journal of Philosophical Research*, Charlottesville, VA., Philosophy Documentation Centre, pp. 23-34.
- Harman, G. and Kulkarni, S. (2007). *Reliable Reasoning*. MIT Press.
- Harman, G. (2009). Field on the Normative Role of Logic. *Proceedings of the Aristotelian Society* 109 (1): 333-335.
- Horwich, P. (1997). Implicit Definition, Analytic Truth, and Apriori Knowledge. *Noûs* 31 (4): 423-440.
- Horwich, P. (1998a). *Truth*. Oxford University Press.
- Horwich, P. (1998b). *Meaning*. Oxford University Press.
- Horwich, P. (2000). Stipulation, Meaning and Apriority. In Boghossian and Peacocke 2000: 150-169.
- Horwich, P. (2005). *Reflections on Meaning*. Oxford University Press.
- Horwich, P. (2008). Ungrounded Reason. *Journal of Philosophy* 105 (9): 453-471.
- Hawthorne, J. (2004). *Knowledge and Lotteries*. Oxford University Press.
- Hawthorne, J. and Stanley, J. (2008). Knowledge and Action. *Journal of Philosophy* 105 (10): 571-590.
- Hume, D. (1888). *A Treatise of Human Nature* (ed. L. A. Selby-Bigge). Oxford University Press, 1978.
- Kaplan, D. What is Meaning? Explorations in the Theory of *Meaning as Use*. Manuscript.
- Katz, J. (2000). *Realistic Rationalism*. MIT Press.
- Kenny, A. (1970). *Descartes: Philosophical Letters*. Oxford University Press.

- Klaczynski, P. A., and Gordon, D. H. (1996). Self-serving influences on adolescents' evaluations of belief-relevant evidence. *Journal of Experimental Child Psychology* 62: 317–339.
- Klaczynski, P. A. (1997). Bias in adolescents' everyday reasoning and its relationship with intellectual ability, personal theories, and self-serving motivation. *Developmental Psychology* 33: 273–283.
- Klaczynski, P. A., Gordon, D. H., and Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology* 89: 470–485.
- Klaczynski, P. A., and Robinson, B. (2000). Personal theories, intellectual ability, and epistemological beliefs: Adult age differences in everyday reasoning tasks. *Psychology and Aging* 15: 400–416.
- Klaczynski, P. A., and Lavalley, K. L. (2005). Domain-specific identity, epistemic regulation, and intellectual ability as predictors of belief-based reasoning: A dual-process perspective. *Journal of Experimental Child Psychology*.
- Kornblith, H. (2000). The Impurity of Reason. *Pacific Philosophical Quarterly* 81: 67-89.
- Kornblith, H. (1998). The Role of Intuition in Philosophical Inquiry. In DePaul and Ramsey 1998: 129-141.
- Koslow, A. (1992). *A Structuralist Theory of Logic*. Cambridge University Press.
- Kripke, S. (1974a). Princeton Lectures on the Nature of Logic. Transcription of tapes. Manuscript.
- Kripke, S. (1974b). The Question of Logic. Transcription of a lecture given at the University of Pittsburg. Manuscript.
- Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Harvard University Press.

- Kripke, S. (2011a). *Philosophical Troubles*, Oxford, Oxford University Press.
- Kripke, S. (2011b). Vacuous Names and Fictional Entities. In Kripke 2011a: 52-74.
- Kripke, S. (2011c). The First Person. In Kripke 2011a: 292-321.
- Kusch, M. (2006). *A Sceptical Guide to Meaning and Rules: Defending Kripke's Wittgenstein*. McGill Queen's University Press.
- Leonardi, P., and Santambrogio, M. (eds.), (1995). *On Quine. New Essays*. Cambridge University Press.
- Levin, M. E. (1978). Quine's View(s) of Logical Truth. *Southwestern Journal of Philosophy* 9 (2): 45-67.
- Lewis, D. (1969). *Convention*. Harvard University Press.
- Lurz, R. (ed.), (2009). *Philosophy of Animal Minds*. Cambridge University Press.
- Lycan, W. (1993). MPP, Rip. In *Philosophical Perspectives, 7, Language and Logic*, Tomberlin, J. (ed.): 411-428.
- MacFarlane, J. (2004). In What Sense, If Anything, is Logic Normative for Thought?
Unpublished manuscript, available at: www.johnmacfarlane.net/normativity_of_logic.pdf
- McGee, V. (1985). A Counterexample to Modus Ponens. *Journal of Philosophy* 82 (9): 462-471.
- Nagel, T. (1997). *The Last Word*. Oxford University Press.
- Noë, A. (2005). Against Intellectualism. *Analysis* 65 (288): 278-290.
- Nozick, R. (1981). *Philosophical Explanations*. Harvard University Press.
- Nozick, R. (2001). *Invariances: The Structure of the Objective World*. Harvard University Press.
- Peacocke, C. (1987). Understanding Logical Constants: A Realist's Account. *Proceedings of the British Academy* 73: 153-200.
- Peacocke, C. (1992). *A Study of Concepts*. MIT Press.

- Peacocke, C. (1993). How Are A Priori Truths Possible. *European Journal of Philosophy* 1 (2): 157-199.
- Peacocke, C. (1998a). Implicit Conception, Understanding and Rationality. *Philosophical Issues* 9: 43-87.
- Peacocke, C. (1998b). Implicit Conception, the A Priori, and the Identity of Concepts. *Philosophical Issues* 9:121-148.
- Peacocke, C. (1999). *Being Known*. Oxford University Press.
- Peacocke, C. (2000). Explaining the A Priori: The Programme of Moderate Rationalism. In Boghossian and Peacocke 2000: 121-148.
- Peacocke, C. (2004). *The Realm of Reason*. Clarendon Press.
- Peacocke, C. (2008). *Truly Understood*. Oxford University Press.
- Peirce, C. S. (1903a). "On Existential Graphs, Euler's Diagrams, and Logical Algebra." In the *Collected Papers of Charles Sanders Peirce. Vol. 4*. Hartshorne, C. and Weiss, P. (eds.), Harvard University Press, 1933.
- Peirce, C. S. (1903b). *Harvard Lectures on Pragmatism: Lecture IV. The Seven Systems of Metaphysics*. In *Pragmatism As A Principle and Method of Right Thinking: The 1903 "Harvard Lectures On Pragmatism."* Edited by Turrisi, P. (ed.), State University of New York Press, 1997.
- Pettit, P. (1990). The reality of rule-following. *Mind* 99 (393): 1-21.
- Pettit, P. (1990). Affirming the reality of rule-following. *Mind* 99 (395): 433-439.
- Pietarinen, A. (2005). Cultivating Habits of Reason: Peirce and the *Logica Utens* versus *Logica Docens* Distinction. *History of Philosophy Quarterly* 22 (4): 357-372.
- Philie, P. (2007). Carroll's Regress and the Epistemology of Logic. *Philosophical Studies* 134: 183-210.

- Plato, (1961), *Meno*, edited by R. S. Bluck, Cambridge, Cambridge University Press.
- Prior, A. (1960). The Runabout Inference-Ticket. In *Philosophical Logic*. P.F. Strawson (ed.), Oxford University Press, 1967.
- Priest, G. (1979). Two Dogmas of Quineanism. *Philosophical Quarterly* 29 (116): 289-301.
- Priest, G. (2003). On Alternative Geometries, Arithmetics, and Logics: A Tribute to Lukasiewicz. *Studia Logica* 74 (3): 441-468.
- Putnam, H. (1962). The Analytic and the Synthetic. In *Midwest Studies in Philosophy*, Feigl, H. and Maxwell, G. (eds.), Minnesota University Press. Reprinted in Putnam 1975b: 33-69.
- Putnam, H. (1968). Is Logic Empirical? In *Boston Studies in the Philosophy of Science* 5, Reidel, 1968. Reprinted as "The Logic of Quantum Mechanics" in Putnam 1975a: 174-197.
- Putnam, H. (1975a). *Realism and Reason: Philosophical Papers, Vol. I*. Cambridge University Press.
- Putnam, H. (1975b). *Mind, Language and Reality: Philosophical Papers, Vol. II*. Cambridge University Press.
- Putnam, H. (1978). There is at least one A Priori Truth. *Erkenntnis* 13: 153-170.
- Quine, W. V. O. (1936). Truth by Convention. In *Philosophical Essays for A. N. Whitehead*. Lee, O. H. (ed.), Longmans: 90-124. Reprinted in Quine 1966: 70-99.
- Quine, W. V. O., (1951). Two Dogmas of Empiricism. *Philosophical Review* 60 (1): 20-43. Reprinted in Quine 1953: 20-46.
- Quine, W. V. O. (1953). *From a Logical Point of View*. Harvard University Press. Revised 2nd edition in 1961.
- Quine, W. V. O. (1954). Carnap and Logical Truth. *Synthese* 12 (4): 350-374. Reprinted in Quine 1966: 100-125.
- Quine, W. V. O. (1960). *Word and Object*. MIT Press.

- Quine, W. V. O. (1966). *The Ways of Paradox and Other Essays*. Random House. Revised and enlarged edition, Harvard University Press, 1976.
- Quine (1969). Preface. In Lewis (1969).
- Quine, W. V. O. and Ullian, J. S. (1970a). *The Web of Belief*. Random House.
- Quine, W. V. O. (1970b). *Philosophy of Logic*. Harvard University Press.
- Quine, W. V. O. (1986). Reply to Vullemin. In *The Philosophy of W.V Quine*,
- Quine, W. V. O. (1991). Two Dogmas in Retrospect. *Canadian Journal of Philosophy* 21 (3): 265-274.
- Quine, W. V. O. (1995). Reactions. In Leonardi and Santambrogio (1995).
- Railton, P. (2000). A Priori Rules: Wittgenstein on the Normativity of Logic. In Boghossian and Peacocke (2000).
- Rayo, A. and Uzquiano, G. (eds.) (2006). *Absolute Generality*. Oxford University Press.
- Reber, A. S. (1989). Implicit Learning and Tacit Knowledge. *Journal of Experimental Psychology: General* 118: 219-35.
- Reber, A. S. (2003). Implicit Learning. In Nadel 2003, vol. 2: 486-91.
- Rees, J. (1951). What Achilles Said to the Tortoise. *Mind* 60 (238): 241-246.
- Rescher, N. (1969). *Many-Valued Logic*. McGraw Hill.
- Rescola, M. (2009). Chrysippus' Dog a Case in Non-Linguistics Cognition. In *Philosophy of Animal Minds*, Lurz, R. (ed.), Cambridge, Cambridge University Press: 52-71.
- Resnik, M. (1996). Quine, the Argument from Proxy Functions, and Structuralism. *Philosophical Topics* 21(1): 129-148.
- Rey, G. (1998). A Naturalistic A Priori. *Philosophical Studies* 92: 25-43.

- Rey, G. (2013). The possibility of a naturalistic Cartesianism regarding intuitions and introspection. In *Philosophical Methodology: The Armchair or the Laboratory?* Haug, M. C. (ed.). Routledge: 243-267.
- Russell, B. (1903). *The Principles of Mathematics*, Vol. 1. Cambridge University Press.
- Russell, B. (1912). *The Problems of Philosophy*. Barnes & Noble Books.
- Russell, G. (2014). Metaphysical analyticity and the epistemology of logic. *Philosophical Studies* 171 (1): 161-175.
- Ryle, G. (1945-46). Knowing How and Knowing That. *Proceedings of the Aristotelian Society* 46: 1-16.
- Ryle, G. (1949). Meaning and Necessity. *Philosophy* 24(88): 69-76.
- Ryle, G. (1950). “If”, “So”, and “Because.” *Philosophical Analysis*, Black, M. (ed), Englewood Cliffs, Prentice-Hall: 302-318. Reprinted in Ryle (1971).
- Ryle, G. (1949). *The Concept of Mind*. Hutchinson and Co.
- Ryle, G. (1971). *Collected Papers*. Hutchinson and Co.
- Sainsbury, R.M. (2002). What logic should we think with? In *Logic, Thought and Language*, O’Hear, A. (ed.). Cambridge University Press: 1-18.
- Schechter, J. and Enoch, D. (2006). Meaning and Justification: The Case of Modus Ponens. *Noûs* 40 (4): 687-715.
- Schechter, J. and Enoch, D. (2008). How Are Basic Belief-Forming Methods Justified. *Philosophy and Phenomenological Research* 76 (3): 547–579.
- Schechter, J. (2013). Could Evolution Explain Our Reliability about Logic? In Tamar Szabo Gendler & John Hawthorne (eds.), *Oxford Studies in Epistemology* 4.
- Schueler, G. (1995). Why “Oughts” are not Facts. *Mind* 104 (416): 712-723.
- Shapiro, S. (2000). The Status of Logic. In Boghossian and Peacocke 2000: 333-366.

- Sher, G. (2010). Epistemic Friction: Reflections on Knowledge, Truth, and Logic. *Erkenntnis* 72(2): 151-176.
- Sher, G. (2011). Is Logic in the Mind or in the World? *Synthese* 181 (2): 353-365.
- Sher, G. (2013). The Foundational Problem of Logic. *Bulletin of Symbolic Logic* 19(2): 145-198.
- Sklar, L. (1974). *Space, Time, and Spacetime*. University of California Press.
- Smiley, T. (1995). A Tale of Two Tortoises. *Mind* 104 (416): 725-736.
- Simchen, O. (2001). Rules and Mentions. *Philosophical Quarterly* 51 (205): 455-473.
- Snowdon, P. (2004). Knowing How and Knowing That: A Distinction Reconsidered. *Proceedings of the Aristotelian Society* 104 (1): 1–29.
- Snowdon, P. F. (2011). Rylean Arguments: Ancient and Modern. In Bengson and Moffett (eds.) 2011.
- Soames, S. (1998). Scepticism about Meaning: Indeterminacy, Normativity, and the Rule-Following Paradox. *Canadian Journal of Philosophy*, suppl. vol. 23: 211-249.
- Sober, E. (2000). Quine. *Proceedings of Aristotelian Society, Supplementary Volume* 74: 237-280.
- Sorabji, R. (1993). *Animal Minds and Human Morals*. Cornell University Press.
- Sperber, D. (1982). Apparently Irrational Beliefs. In *Rationality and Relativism*, Hollis and Lukes (eds.), MIT Press, 1994.
- Stabler, E. (1987). Kripke on Functionalism and Automata. *Synthese* 70: 1-22.
- Stanley, J., and Williamson, T. (2001). Knowing How. *Journal of Philosophy* 98: 411-444.
- Stanley, J. (2005). *Knowledge and Practical Interests*. Oxford University Press.
- Stanley, J. (2011a). *Know How*. Oxford University Press.
- Stanley, J. (2011b). Knowing (How). *Noûs* 45(2): 207-238.

- Stanley, J. and Krakauer, J. W. (2013). Motor skill depends on knowledge of facts. *Frontiers in Human Neuroscience* 7 (503).
- Stanovich, K., and West, R. (2000). Individual Differences in Reasoning: Implications for the Rationality Debate? *Behavioral and Brain Sciences* 23 (5): 645-665.
- Stanovich (2008). What IQ tests miss: The cognitive science of rational and irrational thinking. Yale University Press.
- Steup, M. and Sosa, E. (eds.) (2005). *Contemporary Debates in Epistemology*. Blackwell.
- Stich, S. (1990). *The Fragmentation of Reason*. MIT Press.
- Strawson, P. T. (1952). *Introduction to Logical Theory*. Methuen and Wiley.
- Sun, R., Merrill, E., and Peterson, T. (2001). From Implicit Skills to Explicit Knowledge: A Bottom-Up Model of Skill Learning. *Cognitive Science* 25: 203-44.
- Tennant, N., (1987). *Anti-Realism and Logic: Truth as Eternal*. Oxford University Press.
- Tennant, N. (1997). *The Taming of the True*. Oxford University Press.
- Tennant, N. (2005). "Rule Circularity and the Justification of Deduction." *Philosophical Quarterly*, 55:221: 625-648.
- Thomson, J. F. (1960). What Achilles Should Have Said to the Tortoise. *Ratio* 3: 95-105.
- Tomberlin, J. (1999). *Philosophical Perspectives*, 13, Epistemology, Blackwell.
- van Cleve, J. (1984). "Reliability, Justification, and the Problem of Induction." *Midwest Studies in Philosophy* 9 (1): 555-567.
- Weinberg, J. (2007). How to Challenge Intuitions Empirically Without Risking Skepticism. *Midwest Studies in Philosophy* 31 (1): 318-343.
- Weinberg, J., Gonnerman, C., Buckner, C. and Alexander, J. (2010). Are Philosophers Expert Intuiters? *Philosophical Psychology* 23 (3): 331-355.

- Williamson, T. (2003). Understanding and Inference. *Proceedings of the Aristotelian Society*, Supplementary Volume 77(1): 249-293.
- Williamson, T. (2007). *The Philosophy of Philosophy*. Blackwell.
- Williamson, T. (2011). Reply to Boghossian. *Philosophy and Phenomenological Research* 82 (2): 498-506.
- Williamson, T. (2012). Boghossian and Casalegno on Understanding and Inference. *Dialectica* 66 (2): 237-247.
- Wilson, M. (1995). Animal ideas. *Proceedings and Addresses of the American Philosophical Association* 69: 7-25.
- Wisdom, W. (1974). Lewis Carroll's Infinite Regress. *Mind* 83 (332): 571-573.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Anscombe G. E. M. and Rhees R. (eds.), Blackwell.
- Wittgenstein, L. (1964). *Remarks on the Foundations of Mathematics*. von Wright, G. H., Rhees R., and Anscombe G. E. M. (eds.). Blackwell.
- Wittgenstein, L. (1969). *On Certainty*. Blackwell.
- Wittgenstein, L. (1982). *Wittgenstein's Lectures: Cambridge, 1930–1932*. D. Lee, ed., University of Chicago Press.
- Wittgenstein, L. (1989). *Lectures on the Foundations of Mathematics. Cambridge 1939: from the notes of R.G. Bosanquet, Norman Malcom, Rush Rhees, and Yorick Smythies*. Diamond, C. (ed.), Chicago University Press.
- Wright, C. (1986). Inventing Logical Necessity. In J. Butterfield (ed.), *Language, Mind and Logic*. Cambridge University Press: 187-209.
- Wright, C. (1989). Necessity, Caution and Scepticism. *Proceedings of the Aristotelian Society*, supp.,

- vol. LXIII: 203-238.
- Wright, C. (1989). Wittgenstein's Rule-following Considerations and the Central Project of Theoretical Linguistics. In A. George (ed.), *Reflections on Chomsky*, Oxford and New York: Basil Blackwell: 233–64. Reprinted in Wright (2001b): 170-213.
- Wright, C. (1992). *Truth and Objectivity*. Harvard University Press.
- Wright, C. (2001a). On Basic Logical Knowledge: Reflections on Paul Boghossian's 'How are Objective Epistemic Reasons Possible?' *Philosophical Studies* 106 (1/2): 41-85.
- Wright, C. (2001b). *Rails to Infinity*. Harvard University Press.
- Wright, C. (2004a). Warrant for Nothing: Notes on Epistemic Entitlement. *Proceedings of the Aristotelian Society*, Supplementary Volume 78: 176-212.
- Wright, C. (2004b). Intuition, Entitlement, and the Epistemology of Logical Laws. *Dialectica* 58 (1): 155-175.
- Wright, C. (2007). Rule Following Without Reasons: Wittgenstein's Quietism and the Constitutive Question. *Ratio* 20 (4): 481-502.
- Wright, C. (2009). Internal-External: Doxastic Norms and the Defusing of Sceptical Paradox. *The Journal of Philosophy* 105 (9): 501–517.
- Wright, C. (2012). Replies Part I: The Rule-Following Considerations and the Normativity of Meaning. In *Mind, Meaning, and Knowledge: Themes from the Philosophy of Crispin Wright*. Oxford University Press: 380- 401.
- Wright, C. (2014). Comment on Paul Boghossian: What Is Inference? *Philosophical Studies* 169: 27-37.

